

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 12:17:38 ; Search time 9447.36 Seconds
(without alignments)
11696.741 Million cell updates/sec

Title: US-09-843-159B-1

Perfect score: 3797

Sequence: 1 cttgaagacactggatttc.....ccttgcgtgaaaaaaaaa 3797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_man.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3388.8	89.2	6028	9	AF309033	AF309033 Homo sapi
2	3388.8	89.2	6123	6	AX057579	AX057579 Sequence
3	3387.2	89.2	3815	9	AF329696	AF329696 Homo sapi
4	3387.2	89.2	4127	6	AX062241	AX062241 Sequence
5	3387.2	89.2	4406	6	AX062247	AX062247 Sequence
6	3387.2	89.2	4992	6	AX062275	AX062275 Sequence
7	3387.2	89.2	5002	6	AX062273	AX062273 Sequence
8	3387.2	89.2	5810	9	AF264912	AF264912 Homo sapi
9	3387.2	89.2	6189	9	AF438201	AF438201 Homo sapi
10	3386	89.2	3394	6	AX133694	AX133694 Sequence
11	3382.8	89.1	4512	6	AX076213	AX076213 Sequence
12	3338.4	87.9	4295	9	AF305081	AF305081 Homo sapi
13	3300.4	86.9	3501	9	AF342982	AF342982 Homo sapi
14	3297	86.8	3400	6	AX029397	AX029397 Sequence
15	3296.2	86.8	3508	6	AX062142	AX062142 Sequence
16	3291.8	86.7	3498	6	AX076214	AX076214 Sequence
17	2774.6	73.1	5005	6	AX052530	AX052530 Sequence
18	2439.8	64.3	2971	6	AX062233	AX062233 Sequence
19	2439.8	64.3	3353	6	AX062229	AX062229 Sequence
20	2357.8	62.1	3799	6	AX062231	AX062231 Sequence
21	1922.4	50.6	4130	9	HS804989	AL833676 Homo sapi
22	1813.4	47.8	3984	6	AX062144	AX062144 Sequence
23	1813.4	47.8	4134	6	AX467061	AX467061 Sequence
24	1813.4	47.8	4134	9	AF082556	AF082556 Homo sapi
25	1794.8	47.3	3254	6	AX057580	AX057580 Sequence
26	1794.8	47.3	8901	9	AF082557	AF082557 Homo sapi
27	1449.6	38.2	4491	6	AX467067	AX467067 Sequence
28	1359.4	35.8	1568	6	AX062239	AX062239 Sequence
29	1297.8	34.2	4657	6	AX467069	AX467069 Sequence
30	1172.2	30.9	1242	6	AX052522	AX052522 Sequence
31	1168.4	30.8	3308	6	AX062318	AX062318 Sequence
32	1084.8	28.6	1103	6	AX062300	AX062300 Sequence
33	1076.8	28.4	3066	9	AX023746	AX023746 Homo sapi
34	1008.8	26.6	5482	3	AF132196	AF132196 Drosophil
35	1008.8	26.6	5482	6	AX062279	AX062279 Sequence
36	994	26.2	2409	6	AX429577	AX429577 Sequence
37	990	26.1	1007	6	AX052525	AX052525 Sequence
38	748.6	19.7	48169	2	AC017396	AC017396 Drosophil
39	748.6	19.7	167396	3	AC010579	AC010579 Drosophil
40	748.6	19.7	225478	3	AE003752	AE003752 Drosophil
41	574	15.1	582	6	AX062186	AX062186 Sequence
42	459	12.1	460	6	AX062146	AX062146 Sequence
43	438.8	11.6	142085	9	AL359707	AL359707 Human DNA
44	396.8	10.5	402	6	AX201878	AX201878 Sequence
45	359.6	9.5	1691	6	AX062182	AX062182 Sequence

ALIGNMENTS

RESULT 1
AF309033
LOCUS AF309033 6028 bp mRNA linear PRI 31-JUL-2001
DEFINITION Homo sapiens tankyrase-2 (TNKS-2) mRNA, complete cds.
ACCESSION AF309033
VERSION AF309033.1 GI:15042551
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 6028)
AUTHORS Chi,N.-W. and Lodish,H.F.
TITLE A novel homolog of Tankyrase
JOURNAL Unpublished

Pred. No. is the number of results predicted by chance to have a

Db 1829 GTCTACACCACTTCATTTTCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCT 1888
Qy 1843 ACAGCATGGAGCTGATGTGATGCTAAAGATAAAGAGGCTTCTACTCTTGGCACAATGC 1902
Db 1889 ACAGCATGGAGCTGATGTGATGCTAAAGATAAAGAGGCTTGTACTCTTGGCACAATGC 1948
Qy 1903 ATGTTCTTATGGACATATGAAGTTGAGAACTCTTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1949 ATGTTCTTATGGACATATGAAGTTGAGAACTCTTTGTTAAACATGGAGCAGTAGTTAA 2008
Qy 1963 TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAGAGAAATATGA 2022
Db 2009 TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAGAGAAATATGA 2068
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAGCAAGAGAGGATGAAA 2082
Db 2069 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAGCAAGAGAGGATGAAA 2128
Qy 2083 TACTCCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTCAGAGTCTGCTTAGGGGGA 2142
Db 2129 TACTCCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTCAGAGTCTGCTTAGGGGGA 2188
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGCTTCTTTAGCCAGAGTGAAGAAGTTGCTTC 2202
Db 2189 TGCAGCTTTGCTAGATGCTGCCAAGAGGCTTCTTTAGCCAGAGTGAAGAAGTTGCTTC 2248
Qy 2203 TCTGTGATAATGTAATTTGCCCGCATACCCAGGAGCAGATTTCAACACCTTTACATTTAGC 2262
Db 2249 TCTGTGATAATGTAATTTGCCCGCATACCCAGGAGCAGATTTCAACACCTTTACATTTAGC 2308
Qy 2263 AGCTGGTTATTAATTTAGAACTTGCAGAGTATTTGTTTACACACGGAGCTGATGTA 2322
Db 2309 AGCTGGTTATTAATTTAGAACTTGCAGAGTATTTGTTTACACACGGAGCTGATGTA 2368
Qy 2323 TGCCCAAGACAAAGAGGAGCTTATTCCTTTTACATAATGAGCATCTTACGGCATCTAGA 2382
Db 2369 TGCCCAAGACAAAGAGGAGCTTATTCCTTTTACATAATGAGCATCTTACGGCATCTAGA 2428
Qy 2383 TGTAGCAGCTCTACTTAATAAGTATTAATGATGTGCAATGCCAGGACAAATGGGCTTT 2442
Db 2429 TGTAGCAGCTCTACTTAATAAGTATTAATGATGTGCAATGCCAGGACAAATGGGCTTT 2488
Qy 2443 CACACCTTTGCAGAGCAGCCCAAGAGGAGCAGACAGCTTTGTGCTTTGCTAGC 2502
Db 2489 CACACCTTTGCAGAGCAGCCCAAGAGGAGCAGACAGCTTTGTGCTTTGCTAGC 2548
Qy 2503 CCAATGAGCTGACCCGACTCTTAAATCAGGAAGACAAACACCTTTAGATTTAGTTTC 2562
Db 2549 CCAATGAGCTGACCCGACTCTTAAATCAGGAAGACAAACACCTTTAGATTTAGTTTC 2608
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTGCGCCTCTTG 2622
Db 2609 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTGCGCCTCTTG 2668
Qy 2623 TTACAGCCTCAAGTCTCAATGGTGTGAGAGCCAGAGCCAGCAGATGCTCTCTC 2682
Db 2669 TTACAGCCTCAAGTCTCAATGGTGTGAGAGCCAGAGCCAGCAGATGCTCTCTC 2728
Qy 2683 TTCAGTTCATCTAGCCCATCAAGCCTTTCTGACGCCAGCAGCTTTGACAACTTATCTGG 2742
Db 2729 TTCAGTTCATCTAGCCCATCAAGCCTTTCTGACGCCAGCAGCTTTGACAACTTATCTGG 2788
Qy 2743 GAGTTTTCAGAACTGCTTCTAGTAGTTAGTTCAAGTGAACAGAGGCTGCTTCCAGTTT 2802
Db 2789 GAGTTTTCAGAACTGCTTCTAGTAGTTAGTTCAAGTGAACAGAGGCTGCTTCCAGTTT 2848
Qy 2803 GGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCAATACTCAATTCGTAAAGAACTCTTGG 2862
Db 2849 GGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCAATACTCAATTCGTAAAGAACTCTTGG 2908
Qy 2863 ACTTGAGCACCTTAATGGATATATTTTGAAGAGAACAGATCATCTTTGGATGTATTAGTTGA 2922
Db 2909 ACTTGAGCACCTTAATGGATATATTTTGAAGAGAACAGATCATCTTTGGATGTATTAGTTGA 2968

Qy 2923 GATGGGCAAGAGGAGCTGAAGGAGATGGAATCAATGCTTATGGACATAGGCACAACT 2982
Db 2969 GATGGGCAAGAGGAGCTGAAGGAGATGGAATCAATGCTTATGGACATAGGCACAACT 3028
Qy 2983 AATTAAGAGCTCGAGAGACTTATCTCCGAGCAACAAGGCTCTTAACCCATATTTAACTTT 3042
Db 3029 AATTAAGAGCTCGAGAGACTTATCTCCGAGCAACAAGGCTCTTAACCCATATTTAACTTT 3088
Qy 3043 GAACACCTCTGCTAGTGGAAACAATTTCTTATAGATCTGTCTCTCTGATGATAAAGAGTTTCA 3102
Db 3089 GAACACCTCTGCTAGTGGAAACAATTTCTTATAGATCTGTCTCTCTGATGATAAAGAGTTTCA 3148
Qy 3103 GTCTGTGGAGAGAGAGATGCAAAAGTACAGTTTCTGAGAGCAGAGATGAGGTCATGCAGG 3162
Db 3149 GTCTGTGGAGAGAGAGATGCAAAAGTACAGTTTCTGAGAGCAGAGATGAGGTCATGCAGG 3208
Qy 3163 TGGAACTCTTCAACAGATACAATATTTCTCAAGATTCAGAAAGTTTCTAACAAGAAACTATG 3222
Db 3209 TGGAACTCTTCAACAGATACAATATTTCTCAAGATTCAGAAAGTTTCTAACAAGAAACTATG 3268
Qy 3223 GGAAAGATACACTCACCGGAGAGAAAGAGTTTCTGAAAGAAACCAACCACTGCCAATGA 3282
Db 3269 GGAAAGATACACTCACCGGAGAGAAAGAGTTTCTGAAAGAAACCAACCACTGCCAATGA 3328
Qy 3283 ACGAATGCTATTTTCAATGCTGCTCTCTTTTGTGAATGCAATTTATCCACAAGGCTTTGATGA 3342
Db 3329 ACGAATGCTATTTTCAATGCTGCTCTCTTTTGTGAATGCAATTTATCCACAAGGCTTTGATGA 3388
Qy 3343 AAGGATCGGTACATAGTGGTGTATTTTGGAGCTGGCAATTTATTTGCTGAAACTCTTC 3402
Db 3389 AAGGATCGGTACATAGTGGTGTATTTTGGAGCTGGCAATTTATTTGCTGAAACTCTTC 3448
Qy 3403 CAAAAGCAATCAATATGATATGGAATTTGGAGAGGTTACTGGTGTCCAGTTTCAACAAGA 3462
Db 3449 CAAAAGCAATCAATATGATATGGAATTTGGAGAGGTTACTGGTGTCCAGTTTCAACAAGA 3508
Qy 3463 CAGATCTTTTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACTTTGGGAAAGTC 3522
Db 3509 CAGATCTTTTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACTTTGGGAAAGTC 3568
Qy 3523 TTTCTCTCAGTTCAGTGCATGCAATGAAATGGCAGATTTCTCTCCAGTCACTCACTCAGTCAC 3582
Db 3569 TTTCTCTCAGTTCAGTGCATGCAATGAAATGGCAGATTTCTCTCCAGTCACTCACTCAGTCAC 3628
Qy 3583 TGGTAGGCCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3642
Db 3629 TGGTAGGCCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3688
Qy 3643 GGCTATCTCAGTATTTTAAATTTACTTACAGATTTATGAGGCTGAAGGCTATGGTTCGATGG 3702
Db 3689 GGCTATCTCAGTATTTTAAATTTACTTACAGATTTATGAGGCTGAAGGCTATGGTTCGATGG 3748
Qy 3703 ATAAATAGTTATTTTAAAGAACTAATTTCCACTGAACCTAAAATCATCAAGCAGCAGTGG 3762
Db 3749 ATAAATAGTTATTTTAAAGAACTAATTTCCACTGAACCTAAAATCATCAAGCAGCAGTGG 3808
Qy 3763 CCTCTAGCTTTTACTCTCTCTGCTCAAAAAAAA 3794
Db 3809 CCTCTAGCTTTTACTCTCTCTGCTGAAAAAAA 3840

RESULT 2
AX057579
LOCUS
DEFINITION
Accession
Version
Keywords
Source
ORGANISM

AX057579 Sequence 2 from Patent WO0077225.
AX057579
AX057579.1 GI:12310301

linear

DNA

6123 bp

PAT 17-JAN-2001

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 6123)
AUTHORS Chi.N.W. and Lodish.H.F.
TITLE A novel insulin signaling molecule
JOURNAL Patent: WO 0077225-A 2 21-DEC-2000;
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) : THE GENERAL
HOSPITAL CORPORATION (US)
FEATURES Location/Qualifiers
source
1..6123
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1758 a 1189 c 1358 g 1818 t
ORIGIN
Query Match 89.2%; Score 3388.8; DB 6; Length 6123;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 403 AGTTTGGCGGAAGAGCTAGTTCAATATTTGCTTCAGATGGTGCAGGTGTCGAAGC 462
DB 449 AGTTTGGCGGAAGAGAGCTAGTTCAATATTTGCTTCAGATGGTGCAGGTGTCGAAGC 508
QY 463 ACCTGATGATGGGGCTTATTCCTCTTCATATGATGCTCTTTTGGTCAATGCTGAAGT 522
DB 509 ACCTGATGATGGGGCTTATTCCTCTTCATATGATGCTCTTTTGGTCAATGCTGAAGT 568
QY 523 AGTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATAC 582
DB 569 AGTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATAC 628
QY 583 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTGATTTGCTTGTGCTGTGTACAGCA 642
DB 629 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTGATTTGCTTGTGCTGTGTACAGCA 688
QY 643 TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGACAGCATTTGGATTAGCAGATCC 702
DB 689 TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGACAGCATTTGGATTAGCAGATCC 748
QY 703 ATCTGCAAGAGCTGCTTACTGCTGAATATAGAAAGATGAATCTTACAAAGTGCCAG 762
DB 749 ATCTGCAAGAGCTGCTTACTGCTGAATATAGAAAGATGAATCTTACAAAGTGCCAG 808
QY 763 GAGTGGCAATGAAGAAATGATGGCTTACTCACACCATTTAAATGTCAACTGCCACGC 822
DB 809 GAGTGGCAATGAAGAAATGATGGCTTACTCACACCATTTAAATGTCAACTGCCACGC 868
QY 823 AAGTGATGGCAGAAAGTCAATCCATTACATTTGGCAGCAGGATATACAGAGTAAAGAT 882
DB 869 AAGTGATGGCAGAAAGTCAATCCATTACATTTGGCAGCAGGATATACAGAGTAAAGAT 928
QY 883 TGACAGCTGTTACTGCAACATGGAGTGTGCTCATGCTTAAAGATAAAGTGATCTGGT 942
DB 929 TGACAGCTGTTACTGCAACATGGAGTGTGCTCATGCTTAAAGATAAAGTGATCTGGT 988
QY 943 ACCATTACAAATGCCCTGTTCTTATGGTCAATATGAAGTAACTGAATTTGGTCAAGCA 1002
DB 989 ACCATTACAAATGCCCTGTTCTTATGGTCAATATGAAGTAACTGAATTTGGTCAAGCA 1048
QY 1003 TGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTTCATGAGGAGCTTC 1062
DB 1049 TGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTTCATGAGGAGCTTC 1108
QY 1063 TAAGAACAGGCTTGAAGTATGCTTCTTCTTAAAGTATGGTGCAGACCCCAACACTGCT 1122
DB 1109 TAAGAACAGGCTTGAAGTATGCTTCTTCTTAAAGTATGGTGCAGACCCCAACACTGCT 1168
QY 1123 CAATTGTCACAAATAAAGTGTGTATAGACTTGGTCCCAACACCAAGTTAAAGAAAGATT 1182
DB 1169 CAATTGTCACAAATAAAGTGTGTATAGACTTGGTCCCAACACCAAGTTAAAGAAAGATT 1228
QY 1183 AGCATATGAATTTAAAGGCCACTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTCG 1242
DB 1229 AGCATATGAATTTAAAGGCCACTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTCG 1288

QY 1243 AATCAAAACACATCTCTCTGAAATGCTGAATTTCAAGCATCTTCAACACATGAAC 1302
DB 1289 AATCAAAACACATCTCTCTGAAATGCTGAATTTCAAGCATCTTCAACACATGAAC 1348
QY 1303 AGCATTTGCTTGTGCTGCTGCATTCATATCCCAAAAGAAACAAATATGTGAATGTT 1362
DB 1349 AGCATTTGCTTGTGCTGCTGCATTCATATCCCAAAAGAAACAAATATGTGAATGTT 1408
QY 1363 GCTAAGAAAGAGCAACACATCAATGAAAGACTAAAGATTTCTTGACCTCTGCACGT 1422
DB 1409 GCTAAGAAAGAGCAACACATCAATGAAAGACTAAAGATTTCTTGACCTCTGCACGT 1468
QY 1423 GGCATCTGAGAAAGCTCATATAATGATGTTGTTGAAGTAGTGGTGAACACGAAGGT 1482
DB 1469 GGCATCTGAGAAAGCTCATATAATGATGTTGTTGAAGTAGTGGTGAACACGAAGGT 1528
QY 1483 TAATGCTTGGATAAATCTTGGTGCAGACTTCTTACACAGAGCTGCATATTTGGTCACT 1542
DB 1529 TAATGCTTGGATAAATCTTGGTGCAGACTTCTTACACAGAGCTGCATATTTGGTCACT 1588
QY 1543 ACAAACTGCCGCTACTCTGAGCTATGGTGTGATCTTAACATATATCCCTTCAGGG 1602
DB 1589 ACAAACTGCCGCTACTCTGAGCTATGGTGTGATCTTAACATATATCCCTTCAGGG 1648
QY 1603 CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCCACAGAGGTATCTC 1662
DB 1649 CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCCACAGAGGTATCTC 1708
QY 1663 ATTAGGTAATTCAGAGCAGACAGACAATTCCTTGGAGCTGCAAGAGCTGGAGATGCGA 1722
DB 1709 ATTAGGTAATTCAGAGCAGACAGACAATTCCTTGGAGCTGCAAGAGCTGGAGATGCGA 1768
QY 1723 AACTGTAAAAAATCTGTTACTGTTGAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCA 1782
DB 1769 AACTGTAAAAAATCTGTTACTGTTGAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCA 1828
QY 1783 GTCTACACCACCTTCATTTTCAGCTGGGTATACAGAGTGTCCGTGTGGAATATCTGCT 1842
DB 1829 GTCTACACCACCTTCATTTTCAGCTGGGTATACAGAGTGTCCGTGTGGAATATCTGCT 1888
QY 1843 ACAGCATGGAGCTGATGCTGATGCTAAAGATAAGAGAGGCTTGTACCTTTGCAACAATGC 1902
DB 1889 ACAGCATGGAGCTGATGCTGATGCTAAAGATAAGAGAGGCTTGTACCTTTGCAACAATGC 1948
QY 1903 ATGTTCTTATGGACATTTATGAAGTTGCAGAACTTCTTTTAAACATGAGCAGTAGTTAA 1962
DB 1949 ATGTTCTTATGGACATTTATGAAGTTGCAGAACTTCTTTTAAACATGAGCAGTAGTTAA 2008
QY 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAAGCAGCAGCAAAAGAAATATGA 2022
DB 2009 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAAGCAGCAGCAAAAGAAATATGA 2068
QY 2023 AATTGTCAAACTCTGCTCCAGCATGTTGCAAGCCCTACCAAAAAACAGGGATGAAA 2082
DB 2069 AATTGTCAAACTCTGCTCCAGCATGTTGCAAGCCCTACCAAAAAACAGGGATGAAA 2128
QY 2083 TACTCTTTTGGATCTTGTAAAGATGGAGATACAGATATTCAGATCTGCTTTAGGGGAGA 2142
DB 2129 TACTCTTTTGGATCTTGTAAAGATGGAGATACAGATATTCAGATCTGCTTTAGGGGAGA 2188
QY 2143 TGCAGCTTGTGATGCTGCCAAGAGGTTCTTTAGCCAGAGTGAAGAAAGTTGCTTTC 2202
DB 2189 TGCAGCTTGTGATGCTGCCAAGAGGTTGTTTACCCAGAGTGAAGAAAGTTGCTTTC 2248
QY 2203 TCCTGATAATGATAATTCGCCGCTATACCAAGGAGACATTCACACCTTTTACATTTAGC 2262
DB 2249 TCCTGATAATGATAATTCGCCGCTATACCAAGGAGACATTCACACCTTTTACATTTAGC 2308
QY 2263 AGCTGGTTATATTAATTTAGAGTTGAGAGTATTTGTTACACAGCAGCTGATGTTGAA 2322
DB 2309 AGCTGGTTATATTAATTTAGAGTTGAGAGTATTTGTTTACACAGCAGCTGATGTTGAA 2368
QY 2323 TGCCCAACAGCAAGAGGAGGACTTTATTCCTTTTACATAATCAGCATCTTACGGGCGATGAGA 2382

CSYGHVEAELLVKGVVNVADLWKFTPLHEAAAKGKYEICLLQLHQHGDPTKKNRD
GNTPLDVLKGDGTDIDQLLRGDAALDAAKGLCLARVKKLSPDNVNCRDTOGRHSTP
LHLAAGYNLEAYEYLLOHGADVNAQDGLPLHNAASYGHVDVAALLIKYNACVNA
TDKWAFTPLHEAAQKGRTOCALLAHAGDPTLKNOEGOTPLDLVSADVDVSAALLTAAM
PPSALPSCVKPOVLNCRSPGATADALSSGSPSSLSAASSLDNLNLSFSELSSVVS
SSGTGASLEKEVPGVDFSTIQVRNLGLEHLMDFEREOITLDLIVEMGHKEIKE
IGINAYGHRHKLKIGVERLISGOQGLNPLYLLNTSGSGTILLDLSPDDKEFOSVEEM
QSTVREHDDGHAGIFRNRYLTKIOKVCNKKLWERYTHRRKEVEENHNHANERMLF
HGSFPVNAITHKGFDERHAYIGCMFCAGIYFAENSXSNOYVYIGGGTGCPVHKDRS
CVICHROLFCRVTLCKSPLOFSAMKMAHSPGPHSVTGRPSVNGLALAEYVYIRGEQ
AYPEYLTVQINRPSGMDG"

BASE COUNT 1120 a 788 c 924 g 983 t
ORIGIN

Query Match 89.2%; Score 3387.2; DB 9; Length 3815;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	403	AGGTTTGGCGGAAAGACGTAGTTGAATAATTTGCTTCAGAAATGGTGCAAGTGTCCAAGC	462
Db	337	AGGTTTGGCGGAAAGACGTAGTTGAATAATTTGCTTCAGAAATGGTGCAAAATGCCAAGC	396
QY	463	ACGTGATGATGGGGCTTATTCCTCTTCATAATGCATGCTCTTTGGTCATGCTGAAGT	522
Db	397	ACGTGATGATGGGGCTTATTCCTCTTCATAATGCATGCTCTTTGGTCATGCTGAAGT	456
QY	523	AGTCAATCTCTTTTTCGCACATGGTGACAGACCCCAATGCTCGAGATAATTTGGAATTATAC	582
Db	457	AGTCAATCTCTTTTTCGCACATGGTGACAGACCCCAATGCTCGAGATAATTTGGAATTATAC	516
QY	583	TCCTCTCCATGAAGCTTGAATTAAGGAAGATTTGATTTGGTCTGTGTACAGCA	642
Db	517	TCCTCTCCATGAAGCTTGAATTAAGGAAGATTTGATTTGGTCTGTGTACAGCA	576
QY	643	TGGAGCTGAGCCCAACCATCCGAATACAGATGAAGGACAGCATTTGGATTTAGCAGATCC	702
Db	577	TGGAGCTGAGCCCAACCATCCGAATACAGATGAAGGACAGCATTTGGATTTAGCAGATCC	636
QY	703	ATCTGCCAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAATCTTTAGAAAGTCCAG	762
Db	637	ATCTGCCAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAATCTTTAGAAAGTCCAG	696
QY	763	GAGTGGCAATGAAGAAATGATGCTCTACTCACACCATTAATTAATGTCACCTGCCACGC	822
Db	697	GAGTGGCAATGAAGAAATGATGCTCTACTCACACCATTAATTAATGTCACCTGCCACGC	756
QY	823	AGTGTGCGAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT	882
Db	757	AGTGTGCGAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT	816
QY	883	TGTACAGCTGTTACTCCACATGGAGCTGATGTCATGCTTAAGATTAAGAGTATCGGT	942
Db	817	TGTACAGCTGTTACTCCACATGGAGCTGATGTCATGCTTAAGATTAAGAGTATCGGT	876
QY	943	ACCATTAACAATGCCCTGTCTTATGGTTCATTAAGTAAGTAACTGAACTTTTGGTCAAGCA	1002
Db	877	ACCATTAACAATGCCCTGTCTTATGGTTCATTAAGTAAGTAACTGAACTTTTGGTCAAGCA	936
QY	1003	TGGTGCTGTGTAATGCAATGGAGCTTGTGGCAATTCACCTCTTTCATCAGGACGCTTC	1062
Db	937	TGGTGCTGTGTAATGCAATGGAGCTTGTGGCAATTCACCTCTTTCATCAGGACGCTTC	996
QY	1063	TAAAGACAGGTTGAAGTATGTTCTCTTCTTAAGTTATGGTGCAGACCCCAACACTGCT	1122
Db	997	TAAAGACAGGTTGAAGTATGTTCTCTTCTTAAGTTATGGTGCAGACCCCAACACTGCT	1056
QY	1123	CAATTGTCACAAATAAAGTGCTATAGACTTGGCTCCACACACACAGTTTAAAGAAAGATT	1182
Db	1057	CAATTGTCACAAATAAAGTGCTATAGACTTGGCTCCACACACACAGTTTAAAGAAAGATT	1116
QY	1183	AGCATATGAATTTAAAGGCCACTTCGTTGCTGCAAGCTGCACGAGAGCTGATGTTACTCG	1242
Db	1117	AGCATATGAATTTAAAGGCCACTTCGTTGCTGCAAGCTGCACGAGAGCTGATGTTACTCG	1176

QY	1243	AATCAAAAACATCTCTCTCTGGAAATGGTGAATTTTCAAGCATCTCTCAAAACACATGAAC	1302
Db	1177	AATCAAAAACATCTCTCTCTGGAAATGGTGAATTTTCAAGCATCTCTCAAAACACATGAAC	1236
QY	1303	AGCATTTGCATTTGCTGCTGCATCTCCATATCCCAAAAGAAACAAATATGTGAACATGTT	1362
Db	1237	AGCATTTGCATTTGCTGCTGCATCTCCATATCCCAAAAGAAACAAATATGTGAACATGTT	1296
QY	1363	GCTAAGAAAGAGACCAACATCAATGAAAAGACTTAAAGAAATCTTGACTCTCTGCAAGT	1422
Db	1297	GCTAAGAAAGAGACCAACATCAATGAAAAGACTTAAAGAAATCTTGACTCTCTGCAAGT	1356
QY	1423	GGCATCTGAGAAAAGCTCATTAATGATTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGT	1482
Db	1357	GGCATCTGAGAAAAGCTCATTAATGATTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGT	1416
QY	1483	TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCAATATTTGGTCATCT	1542
Db	1417	TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCAATATTTGGTCATCT	1476
QY	1543	ACAAACCTGCGCCTACTCTGAGCTATGGTGTGAATTTGCTTAACATATATCCCTTCAGGG	1602
Db	1477	ACAAACCTGCGCCTACTCTGAGCTATGGTGTGAATTTGCTTAACATATATCCCTTCAGGG	1536
QY	1603	CTTTACTGCTTTTACAGATGGAAATGAAATGTACAGCAACTCTCTCCAAGAGGATATCTC	1662
Db	1537	CTTTACTGCTTTTACAGATGGAAATGAAATGTACAGCAACTCTCTCCAAGAGGATATCTC	1596
QY	1663	ATTAGTAATTCAGAGGCACAGACAGAAATTTGCTGGAAGCTGCAAGAGGCTGCA	1722
Db	1597	ATTAGTAATTCAGAGGCACAGACAGAAATTTGCTGGAAGCTGCAAGAGGCTGCA	1656
QY	1723	AACGTAAAAAAACCTGTTACTGTTTACAGAGTGTCAACTGCAGAGACATTTCAAGGGGTCA	1782
Db	1657	AACGTAAAAAAACCTGTTACTGTTTACAGAGTGTCAACTGCAGAGACATTTCAAGGGGTCA	1716
QY	1783	GTCTACACACTTCAATTTTGCAGCTGGGTATAACAGAGTGTGCGTGGTGAATATCTGCT	1842
Db	1717	GTCTACACACTTCAATTTTGCAGCTGGGTATAACAGAGTGTGCGTGGTGAATATCTGCT	1776
QY	1843	ACAGCATGGAGCTGATGTCATGCTTAAAGATAAGAGAGCCTTGTACCTTTGCAACATGC	1902
Db	1777	ACAGCATGGAGCTGATGTCATGCTTAAAGATAAGAGAGCCTTGTACCTTTGCAACATGC	1836
QY	1903	ATGTTCTTATGACATTTAAGTTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTAA	1962
Db	1837	ATGTTCTTATGACATTTAAGTTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTAA	1896
QY	1963	TGTAGCTGATTTATGAAATTTTACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGA	2022
Db	1897	TGTAGCTGATTTATGAAATTTTACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGA	1956
QY	2023	AATTTGCAAACTTCTGCTCCAGCATGTCAGACCTTACCACAAAAAAGAGGATGAAA	2082
Db	1957	AATTTGCAAACTTCTGCTCCAGCATGTCAGACCTTACCACAAAAAAGAGGATGAAA	2016
QY	2083	TACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGACTCTGCTTAGGGGAGA	2142
Db	2017	TACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGACTCTGCTTAGGGGAGA	2076
QY	2143	TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTTAGCCAGAGTGAAGAGTTGCTTTC	2202
Db	2077	TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTTAGCCAGAGTGAAGAGTTGCTTTC	2136
QY	2203	TCTGTATATGTAATTTGCCGGGATACCCAAAGCAGACATTTCAACACCTTTTACATTTAGC	2262
Db	2137	TCTGTATATGTAATTTGCCGGGATACCCAAAGCAGACATTTCAACACCTTTTACATTTAGC	2196
QY	2263	AGCTGTTTATAATTAATTTAGAGTTGCAGAGTATTTTGTACACACGAGCTGATGTGAA	2322
Db	2197	AGCTGTTTATAATTAATTTAGAGTTGCAGAGTATTTTGTACACACGAGCTGATGTGAA	2256

QY	2323	TGCCCCAAGACAAAGAGGAGCTTATTCCTTTACATAATGACAGCATCTTACGGGCGATGTAGA	2382	Db	3337	CAAAAGCAATCAATATGTATATGGAATTCGAGGAGGTACTGGGTGCCAGTTCCACAAGA	3396
Db	2257	TGCCCCAAGACAAAGAGGAGCTTATTCCTTTACATAATGACAGCATCTTACGGGCGATGTAGA	2316	QY	3463	CAGATCTTGTACATTTGCGACAGGAGCTGCTCTTTTCCGGGTAACTTCCGGAAGTC	3522
QY	2383	TGTAGCAGCTCTACTATAAAGATATATGATGTGTCAATGCCACGACAAATGGGCTTT	2442	Db	3397	CAGATCTTGTACATTTGCGACAGGAGCTGCTCTTTTCCGGGTAACTTCCGGAAGTC	3456
Db	2317	TGTAGCAGCTCTACTATAAAGATATATGATGTGTCAATGCCACGACAAATGGGCTTT	2376	QY	3523	TTTCCTGCGAGTCAGTGCATGAAATGACATTCCTCCAGGTCATCCTCAGTCAC	3582
QY	2443	CACACCTTTTGACGAAGCAGCCCAAGGAGGACGAACACAGCTTTGTCTTTGTGCTAGC	2502	Db	3457	TTTCCTGCGAGTCAGTGCATGAAATGACATTCCTCCAGGTCATCCTCAGTCAC	3516
Db	2377	CACACCTTTTGACGAAGCAGCCCAAGGAGGACGAACACAGCTTTGTCTTTGTGCTAGC	2436	QY	3583	TGCTAGGCCAGTGTAAATGGCTAGCATTTAGCTGAATATGTTTATACAGAGGAGACA	3642
QY	2503	CCATGGAGCTGACCCGACTCTTAAATACAGGAAGGACAAACACCTTTAGATTTAGTTTC	2562	Db	3517	TGCTAGGCCAGTGTAAATGGCTAGCATTTAGCTGAATATGTTTATACAGAGGAGACA	3576
Db	2437	CCATGGAGCTGACCCGACTCTTAAATACAGGAAGGACAAACACCTTTAGATTTAGTTTC	2496	QY	3643	GGCTTATCTGAGTATTTAATTAATTTACAGATTTAGAGCCTGAGGTATGCTGATGG	3702
QY	2563	AGCCGATGATGTCAGCGCTCTTCGACAGCAGCCATGCCCCCATCTGCTCGCCCTCTTG	2622	Db	3577	GGCTTATCTGAGTATTTAATTAATTTACAGATTTAGAGCCTGAGGTATGCTGATGG	3636
Db	2497	AGCAGATGATGTCAGCGCTCTTCGACAGCAGCCATGCCCCCATCTGCTCGCCCTCTTG	2556	QY	3703	ATAAATAGTTATTTAAGAACTAATTCACCTGAACCTAAAATCATCAAGCAGCAGTGG	3762
QY	2623	TTACAAGCCTCAAGTGTCTCAATGGTGTGAGAAGCCAGGACCCACTGCAGATGCTCTCTC	2682	Db	3637	ATAAATAGTTATTTAAGAACTAATTCACCTGAACCTAAAATCATCAAGCAGCAGTGG	3696
Db	2557	TTACAAGCCTCAAGTGTCTCAATGGTGTGAGAAGCCAGGACCCACTGCAGATGCTCTCTC	2616	QY	3763	CCTCTACGTTTTCCTCTTCTGCTGAAAAAAA	3794
QY	2683	TTCAAGTCCATCTAGCCCATCAAGCCTTTCTGAGCCAGCAGCTTTGACAACTTATCTGG	2742	Db	3697	CCTCTACGTTTTCCTCTTCTGCTGAAAAAAA	3728
Db	2617	TTCAAGTCCATCTAGCCCATCAAGCCTTTCTGAGCCAGCAGCTTTGACAACTTATCTGG	2676	RESULT 4			
QY	2743	GAGTTTTTCAGAACTGTCTCAGTGTAGTTCAAGTGAAGGACGAGGCTTCCAGTTT	2802	AX062241			
Db	2677	GAGTTTTTCAGAACTGTCTCAGTGTAGTTCAAGTGAAGGACGAGGCTTCCAGTTT	2736	LOCUS			
QY	2803	GGAGAAAGAGGAGTCCAGGAGTAGATTTTATGATCAATTCGTAAGGAATCTTGG	2862	DEFINITION			
Db	2737	GGAGAAAGAGGAGTCCAGGAGTAGATTTTATGATCAATTCGTAAGGAATCTTGG	2796	AX062241			
QY	2863	ACTTGAGCCTCAATGGATATTTTACAGAGACAGATCACTTTGGATGTATAGTTGA	2922	SEQUENCE 100 FROM PATENT WO0100849.			
Db	2797	ACTTGAGCCTCAATGGATATTTTACAGAGACAGATCACTTTGGATGTATAGTTGA	2856	AX062241.1			
QY	2923	GATGGGCAAGAGGAGTGAAGGAGATTTGAAATCAATGCTTATGACATAGGCACAACT	2982	GI:12540142			
Db	2857	GATGGGCAAGAGGAGTGAAGGAGATTTGAAATCAATGCTTATGACATAGGCACAACT	2916	human.			
QY	2983	AATTAAGGAGTCGAGAGCTTATCTCCGACACAAAGCTTTAACCCTTTTAACTTT	3042	Homo sapiens			
Db	2917	AATTAAGGAGTCGAGAGCTTATCTCCGACACAAAGCTTTAACCCTTTTAACTTT	2976	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
QY	3043	GAACACCTCTGGTAGTGGAAACAATTTTATAGATCTGCTCCTGATGATAAGAGTTTCA	3102	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
Db	2977	GAACACCTCTGGTAGTGGAAACAATTTTATAGATCTGCTCCTGATGATAAGAGTTTCA	3036	1 (bases 1 to 4127)			
QY	3103	GTCTGGAGGAGAGATGCAAGTACAGTTTCAGAGACACAGATGGAGTTCATGCAGG	3162	Christenson, E., Demaggio, A.J., Goldman, P.S. and Mcelligott, D.L.			
Db	3037	GTCTGGAGGAGAGATGCAAGTACAGTTTCAGAGACACAGATGGAGTTCATGCAGG	3096	Tankyrase2 materials and methods			
QY	3163	TGGAATCTTCAACAGATCAATTTCTCAAGATTCAGAGGTTTGTACAGAAACTATG	3222	Patent: WO 0100849-A 100 04-JAN-2001;			
Db	3097	TGGAATCTTCAACAGATCAATTTCTCAAGATTCAGAGGTTTGTACAGAAACTATG	3156	ICOS CORPORATION (US)			
QY	3223	GGAAAGATACACTCCCGGAGAGAAAGAGTTTCTCAAGAAACACCAACCATGCCAATGA	3282	Location/Qualifiers			
Db	3157	GGAAAGATACACTCCCGGAGAGAAAGAGTTTCTCAAGAAACACCAACCATGCCAATGA	3216	1..4127			
QY	3283	ACGAATGCTATTTCAATGGTCTCTCTTTTGTGAATGCAATTTCCACAAGCTTTGATGA	3342	/organism="Homo sapiens"			
Db	3217	ACGAATGCTATTTCAATGGTCTCTCTTTTGTGAATGCAATTTCCACAAGCTTTGATGA	3276	/db_xref="taxon:9606"			
QY	3343	AAGCATGCTACATAGTGTATGTTTGGAGCTGGCATTTTATTTGCTGAAAACCTCTTC	3402	<2..3511			
Db	3277	AAGCATGCTACATAGTGTATGTTTGGAGCTGGCATTTTATTTGCTGAAAACCTCTTC	3336	/note="unnamed protein product"			
QY	3403	CAAAAGCAATCAATATGTATGGAATTTGGAGGAGTACTGGGTGCTCCAGTTCCACAAGA	3462	/codon_start=1			
				/protein_id="CAG25060.1"			
				/translation="ARTMSRCAGGAGGACASAAAEVPAARELFEACRNGDVERVK			
				RLVTPKVNSTRDTAGRKSTPLHFAAGFGKDVVEYLQNGANNVQAEEDGGLPLHNAC			
				SFGHAEVNNLLRHGADPNARDNNWYTPLEHAAIKGIDVCIVLQHGAEPTIRNDG			
				RTALDLADPSAKAVLTGEYKKDELLESARGNEEKMMALLTPLNVNCHDGRKSTPL			
				HLAAGYNRKIVQLLQHGADHAKGDPLVHNACSYGHVEITLVLKHCAGYCNAM			
				DLWQFTPLHFAASKNRVEVCSLLLSYGADPTLLNCHNSAIDLAIPPTPKRITCELLRK			
				GSLQQAAREADVTRIKHLSLEMFVNFKHPQTHETALHCAASAPYKPKITCELLRK			
				GANINKEFTPLHFAVEKAHNDVVEVVKHAKVNDLNGLGOTSLHRAAYCGHLQ			
				TCRLILSYGCDPNIISLOQFTALONGENNVQQLQGLISLGNSEADROLLEAAKAGDV			
				ETVKLCTVQSCNDRIEGRQSTPLHFAAGYNRVSVVEYLQHGADHAKGDGLVPL			
				HNACSYGHEVAELLVKHAAGVNVADLWKTPLHFAAAGKYEICLLQHGADPTKK			
				NRDNTPLDLVQDGTIDQDLLRGDAALDAAKGCILARVKKLSSPDNVCNDRDTPK			
				STPLHUAAGYNLLEAVFYLLQHGADNQAQDKGLIPLHNAASYGHVDVAALLIKYNAC			
				VNATDKWFTPLHFAAQKGTQICALLIAGHADPTLKNQEGOTPLDLVSADVDYALLT			
				AMPPSALPSCYKPOVLNVRSPGADATLSSGSPSSLSAASLDNLNLSGFSSELSS			
				VYSSGTEGASLSEKKEVPQVDFSIQFVFNGLGLHLMIDIFEREQITLDLVLEMHGKE			
				LKEIGINAYGHRHKLKGVRLISGQGNLXPLTNTSGSTLTLTDLSPDKFQSVSE			
				EMQSTVREHROGHAGGLFNRYNLIKOKVCKKWLERYTHRRKEVSEENHNANER			
				MIFHSGSPFVNAIHHKGFDERHAYIGMFGAGTYFAENSKSNQYVYGGGTGCPVHK			
				DRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPPHSHSVTGRPSVNGLALAEYVIR			

|||||
Db 2248 CACACCTTTGCAGAACACCCAAAAGGACCAACACAGCTTTGTGCTTTGTGCTAGC 2307
QY 2503 CCATGGAGCTGACCCGACTCTTAAAATCAGGAAGACAAACACCTTTAGATTAGTTTC 2562
Db 2308 CCATGGAGCTGACCCGACTCTTAAAATCAGGAAGACAAACACCTTTAGATTAGTTTC 2367
QY 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGAGCCATGCGCCCATCTGCTGCGCCCTTTG 2622
Db 2368 AGCAGATGATGTCAGCGCTCTTCTGACAGAGCCATGCGCCCATCTGCTGCGCCCTTTG 2427
QY 2623 TTACAAGCCTCAAGTGTCTCAATGCTGTGAGAGCCGAGAGCCACTGCAGATGCTCTCTC 2682
Db 2428 TTACAAGCCTCAAGTGTCTCAATGCTGTGAGAGCCGAGAGCCACTGCAGATGCTCTCTC 2487
QY 2683 TTCAAGTCCATCTAGCCCATCAAGCCCTTTCTGAGAGCCAGAGCTTTTGACAACCTTATCTGG 2742
Db 2488 TTCAAGTCCATCTAGCCCATCAAGCCCTTTCTGAGAGCCAGAGCTTTTGACAACCTTATCTGG 2547
QY 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGGTCTTCCAGTTT 2802
Db 2548 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGGTCTTCCAGTTT 2607
QY 2803 GGAGAAAAGGAGGTTCAGGAGTAGATTTTACGATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2608 GGAGAAAAGGAGGTTCAGGAGTAGATTTTACGATAACTCAATTCGTAAGGAATCTTGG 2667
QY 2863 ACTTGAGCACTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2668 ACTTGAGCACTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2727
QY 2923 GATGGGCAACAGGAGCTGAAGAGATTTGGAATCAATCTTATGACATAGGACAAACT 2982
Db 2728 GATGGGCAACAGGAGCTGAAGAGATTTGGAATCAATCTTATGACATAGGACAAACT 2787
QY 2983 AATTAAGGAGTCCAGAGACTTATCTCCGACAAAGGCTTAAACCATATTAACTTT 3042
Db 2788 AATTAAGGAGTCCAGAGACTTATCTCCGACAAAGGCTTAAACCATATTAACTTT 2847
QY 3043 GAACACCTCTGGTAGTGAACAATTTTATAGATCTGCTCTCTGATGATPAAAGAGTTTCA 3102
Db 2848 GAACACCTCTGGTAGTGAACAATTTTATAGATCTGCTCTCTGATGATPAAAGAGTTTCA 2907
QY 3103 GTCTGTGGAGGAGATGCAAGTACAGTTCTGAGAGCACAGATGAGAGTTCATGCAGG 3162
Db 2908 GTCTGTGGAGGAGATGCAAGTACAGTTCTGAGAGCACAGATGAGAGTTCATGCAGG 2967
QY 3163 TGGAACTCTTCAACAGATACAATATTCTCAAGATTGAGAGGTTTCTGAACAAGAACTATG 3222
Db 2968 TGGAACTCTTCAACAGATACAATATTCTCAAGATTGAGAGGTTTCTGAACAAGAACTATG 3027
QY 3223 GGAAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCAATGCCAATGA 3282
Db 3028 GGAAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCAATGCCAATGA 3087
QY 3283 AGCAATGCTATTTCATGGGCTCTCTTTGTAATGCAATATTCACAAGGCTTTGATGA 3342
Db 3088 AGCAATGCTATTTCATGGGCTCTCTTTGTAATGCAATATTCACAAGGCTTTGATGA 3147
QY 3343 AAGGATGCTGATACAGTGGTATGTTTGGAGCTGGCATTTATTTGCTGAAACCTCTTC 3402
Db 3148 AAGGATGCTGATACAGTGGTATGTTTGGAGCTGGCATTTATTTGCTGAAACCTCTTC 3207
QY 3403 CAAAAGCAATCAATATGATATGGAATTTGAGGAGGTACTGGGTGTCAGGTTTCAAAAAGA 3462
Db 3208 CAAAAGCAATCAATATGATATGGAATTTGAGGAGGTACTGGGTGTCAGGTTTCAAAAAGA 3267
QY 3463 CAGATCTTGTACATTGCGCAGGAGCTGCTCTTTTCCGGGGTAACCTTGGGAAAGTC 3522
Db 3268 CAGATCTTGTACATTGCGCAGGAGCTGCTCTTTTCCGGGGTAACCTTGGGAAAGTC 3327
QY 3523 TTTCCTGCAAGTTTCAAGTGAATGCAATGCAATCTCTCCAGTTCATCACTCAGTCAC 3582
|||||

Db 3328 TTTCCTGCAAGTTTCAAGTGAATGCAATGCAATCTCTCCAGGTCATCACTCAGTCAC 3387
QY 3583 TGGTAGGCCAGTGAATGGCTAGCATTAGCTGAATATGTTATTTTACAGAGAGACA 3642
Db 3388 TGGTAGGCCAGTGAATGGCTAGCATTAGCTGAATATGTTATTTTACAGAGAGACA 3447
QY 3643 GGCCTTATCCTGAGTATTTAAATTTACTTACCAGATTATGAGGCTGAAGTATGGTGCATGG 3702
Db 3448 GGCCTTATCCTGAGTATTTAAATTTACTTACCAGATTATGAGGCTGAAGTATGGTGCATGG 3507
QY 3703 ATAAATAGTATTTTAAAGAACTAATTCCTACTGAACCTTAAATATCAATCAAGCAGCAGTGG 3762
Db 3508 ATAAATAGTATTTTAAAGAACTAATTCCTACTGAACCTTAAATATCAATCAAGCAGCAGTGG 3567
QY 3763 CCTCTACGTTTCTACTCCTTTGCTGAAAAAAA 3794
Db 3568 CCTCTACGTTTCTACTCCTTTGCTGAAAAAAA 3599

RESULT 5
AX062247
LOCUS AX062247 4406 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 106 from Patent WO0100849.
ACCESSION AX062247
VERSION AX062247.1 GI:12540148
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4406)
AUTHORS Christenson, E., Demaggio, A.J., Goldman, P.S. and Mcelligott, D.L.
TITLE Tankyrase2 materials and methods
JOURNAL Patent: WO 0100849-A 106 04-JAN-2001;
ICOS CORPORATION (US)
FEATURES
Location/Qualifiers
1..4406
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS
2..3790
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC25062.1"
/db_xref="GI:12540149"
/translation="ELAGGALPASAAASFODDPGFALPPPPRSGRAGSAPARGRG
GHGTADPPTAGSOAARLASSSPGGLALLAGPGLLLRLALLLALLAARINSGRR
AGGACASAAAEAPARELFACNRGIDVERKLVTPPEKVNSRDTAGRKSTPLHF
AAGGRKDVVEYLLQNGANVOARDGGLIPLHNACSFHAEVNVLLLRHGDNPARDN
WNTPLHEAA IKGIDYIVLLQHGAEPTIRNTDGRATLDADPSKAVLTGYKDE
LLESARGNEKMMALITPLNVNCHASDGRKSTPLAAGYNRVKIVOLLQHGADVH
AKDGLVPLHNACSYHYEVTELLVKHGACVNMIDWOTPLHEAASKNRYEVCSSL
LSYGADPTLLCHNKSAIDLAPTPOLKERLAYEFKSHLSLOAREADYTRIKHLSL
MYNFKHPOTHEALHCNAAAPYKPKQICELLURKGANINEKTEFTPLHVASEKAL
NDVEVVVHEAKNADNLIGQTSLHRAAYCCHLQTCRLLLSYGCDPNIISLQGTAL
OMGENVQQLQECI SLGNSADRLLEAAKAGDVETVKLCYTVSVNCRDIEGROST
PLHFAAGYNRVSYVEYLLQHGADHAKDGLVPLHNACSYHYEVNLLRVHGVVN
VADLWFTPLHEAAAKGKYEICLLQHGADPTKKNRDNTPDLIDVKGDDTDODLLR
GDAALLDAKKGCLARVKLSLSPDNVNCRDQGRHSTPLHLAAGYNLVEAYELLQHG
ADYNAQDKGLIPLHNAASGYHVDVAALLIKYNACVNDKWAFTPLHEAAQGRTOQL
CALLAHAGDPTLKNQEGOTPLDVSADVSAALLTAAMPSPALPSYKPOVLNVRSP
GATADALSSGSPSSPLSLGNSADRLLEAAKAGDVETVKLCYTVSVNCRDIEGROST
STQFVFNILGLEHMDITFEQITLDVLVEMGHKEIGINAYGHRHLIKGVRLI
SGOGLNPYLTNTSGTITLIDLSPPDKFQSFVEEMOSTVREHDDGCHGAGIFNRY
NILKIOKNKKLWERYTHRRKYSNNHANERMLFHGSPFVNAIHKHGFDERHAY
IGMFGAGYFPAENSSKNOYVTYIGGTCGPVHKDRSCYICHRLLFCRVLTGKSLF
QPSAMHAHSPGHHSTGRPSVNGALAEYVIYRGEQAYPEYLYTIQIMRPEGWVGD
"

BASE COUNT 1268 a 910 c 1054 g 1174 t
ORIGIN

Query Match 89.2%; Score 3387.2; DB 6; Length 4406;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	403	AGGTTTTGGCGGAAGACGTAGTTGGAATATTTGCTTCAGAAATGGTGCAAGTGTCCAAGC	462	QY	1483	TAATGCTCGGATAATCTTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATCT	1542
Db	487	AGGTTTTGGCGGAAGACGTAGTTGGAATATTTGCTTCAGAAATGGTGCAAGTGTCCAAGC	546	Db	1567	TAATGCTCGGATAATCTTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATCT	1626
QY	463	AGGTGATGATGGGGCCTTATTCCTCTTCATATATGATGCTCTTTTGGTCAATGCTGAAGT	522	QY	1543	ACAAAGCTCCGCCCTACTCTGAGCTATGGGTGATGCTTAACATTTATCCCTTCAGGG	1602
Db	547	AGGTGATGATGGGGCCTTATTCCTCTTCATATATGATGCTCTTTTGGTCAATGCTGAAGT	606	Db	1627	ACAAAGCTCCGCCCTACTCTGAGCTATGGGTGATGCTTAACATTTATCCCTTCAGGG	1686
QY	523	AGTCAATCTCCTTTTCGGACATGGTGACAGACCCCAATGCTCGAGATAATTTGGAATATATC	582	QY	1603	CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCTCAAGAGGCTATCTC	1662
Db	607	AGTCAATCTCCTTTTCGGACATGGTGACAGACCCCAATGCTCGAGATAATTTGGAATATATC	666	Db	1687	CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCTCAAGAGGCTATCTC	1746
QY	583	TCTCTCCATGAAGCTGCAATTAAGGAAGATTGATGTTTGATTTGCTGTGTACAGCA	642	QY	1663	ATTAGTAATTCAGAGGCACAGACAAATTTGCTGGAAGTGCAGAGCTGCAGATGCGA	1722
Db	667	TCTCTCCATGAAGCTGCAATTAAGGAAGATTGATGTTTGATTTGCTGTGTACAGCA	726	Db	1747	ATTAGTAATTCAGAGGCACAGACAAATTTGCTGGAAGTGCAGAGCTGCAGATGCGA	1806
QY	643	TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGACAGCAATTTGATTTAGCAGATCC	702	QY	1723	AACTGTAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGGTCA	1782
Db	727	TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGACAGCAATTTGATTTAGCAGATCC	786	Db	1807	AACTGTAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGGTCA	1866
QY	703	ATCTGCCAAGCAGTCTTACTGGTGAATATAGAAAGATGAACCTCTTAGAAAGTCCAG	762	QY	1783	GTCTACACCACTTCATTTTCAGCTGGGTATAACAGAGTGTCCGTGTGGAATATCTGCT	1842
Db	787	ATCTGCCAAGCAGTCTTACTGGTGAATATAGAAAGATGAACCTCTTAGAAAGTCCAG	846	Db	1867	GTCTACACCACTTCATTTTCAGCTGGGTATAACAGAGTGTCCGTGTGGAATATCTGCT	1926
QY	763	GAGTGCAATGAAGAAAAATGATGGCTCTACTCACACCATTAATGTTCAACTGCCACGC	822	QY	1843	ACAGCATGGAGCTGATGTGATGCTAAAGATAAAGGAGGCTTTGCTACCTTTGCAATGTC	1902
Db	847	GAGTGCAATGAAGAAAAATGATGGCTCTACTCACACCATTAATGTTCAACTGCCACGC	906	Db	1927	ACAGCATGGAGCTGATGTGATGCTAAAGATAAAGGAGGCTTTGCTACCTTTGCAATGTC	1986
QY	823	AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT	882	QY	1903	ATGTTCTTATGGACATTTACAGTTCAGAACTTCTGTTTAAACATGGAGCAGTAGTTAA	1962
Db	907	AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT	966	Db	1987	ATGTTCTTATGGACATTTACAGTTCAGAACTTCTGTTTAAACATGGAGCAGTAGTTAA	2046
QY	883	TGTACAGCTGTTACTGCACATGAGCTGATGTCATGCTTAAAGATAAAGGATGATCTGGT	942	QY	1963	TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA	2022
Db	967	TGTACAGCTGTTACTGCACATGAGCTGATGTCATGCTTAAAGATAAAGGATGATCTGGT	1026	Db	2047	TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA	2106
QY	943	ACCATTACACAATGCGCTTATGCTCAATTAAGTAACCTGAACCTTTTGGTCAAGCA	1002	QY	2023	AAITTCGAAACTTCTGCTCCAGCATGGTGAGACCCCTACCAAAAAAAGAGGATGAAA	2082
Db	1027	ACCATTACACAATGCGCTTATGCTCAATTAAGTAACCTGAACCTTTTGGTCAAGCA	1086	Db	2107	AAITTCGAAACTTCTGCTCCAGCATGGTGAGACCCCTACCAAAAAAAGAGGATGAAA	2166
QY	1003	TGTTGCTGTGTAATGCAATGAGCTGTGGCAATTCACCTCTTCATGAGGACGCTTC	1062	QY	2083	TACTCTCTTGATCTTGTATAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA	2142
Db	1087	TGTTGCTGTGTAATGCAATGAGCTGTGGCAATTCACCTCTTCATGAGGACGCTTC	1146	Db	2167	TACTCTCTTGATCTTGTATAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA	2226
QY	1063	TAAGAACAGGTTGAAGTATGTTCTCTTCTTAAAGTTATGTTGTCAGACCAACACTGCT	1122	QY	2143	TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGAGTGAAGAAGTTGCTTC	2202
Db	1147	TAAGAACAGGTTGAAGTATGTTCTCTTCTTAAAGTTATGTTGTCAGACCAACACTGCT	1206	Db	2227	TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGAGTGAAGAAGTTGCTTC	2286
QY	1123	CAATTGTCACAATAAAGTGTCTATAGACTTGGCTCCACACCACAGTAAAGAAAGATT	1182	QY	2203	TCCTGATAATGTAATTTGCCGCGATACCCAGGCAGACATTCACACCTTTTACATTTAGC	2262
Db	1207	CAATTGTCACAATAAAGTGTCTATAGACTTGGCTCCACACCACAGTAAAGAAAGATT	1266	Db	2287	TCCTGATAATGTAATTTGCCGCGATACCCAGGCAGACATTCACACCTTTTACATTTAGC	2346
QY	1183	AGCATATGAATTTAAGGCCACTGTTGCTGCAAGCTGCGAGAGAGCTGATGTTACTCG	1242	QY	2263	AGCTGGTTATAAATTTAGAAGTTGCAGAGTATTTGTTTACACACAGCAGAGTGAATGAA	2322
Db	1267	AGCATATGAATTTAAGGCCACTGTTGCTGCAAGCTGCGAGAGAGCTGATGTTACTCG	1326	Db	2347	AGCTGGTTATAAATTTAGAAGTTGCAGAGTATTTGTTTACACACAGCAGAGTGAATGAA	2406
QY	1243	AATCAAAAAACATCTCTCTCTGAAATGTTGAATTTCAAGCATCTCTCAACACATGAAC	1302	QY	2323	TGCCCCAAGACAAAGGAGGACTTATTCCTTTACATAAATGTCAGCATCTTACGGGATGAGA	2382
Db	1327	AATCAAAAAACATCTCTCTCTGAAATGTTGAATTTCAAGCATCTCTCAACACATGAAC	1386	Db	2407	TGCCCCAAGACAAAGGAGGACTTATTCCTTTACATAAATGTCAGCATCTTACGGGATGAGA	2466
QY	1303	AGCATGCAATTTGCTGCTGCATCTCCATATCCCAAGAAAGCAATATGTAACCTGTT	1362	QY	2383	TGTAGCAGCTCTACTATAAAGTATGATGTGTCATGCCACGCAGCAAAATGGGCTTT	2442
Db	1387	AGCATGCAATTTGCTGCTGCATCTCCATATCCCAAGAAAGCAATATGTAACCTGTT	1446	Db	2467	TGTAGCAGCTCTACTATAAAGTATGATGTGTCATGCCACGCAGCAAAATGGGCTTT	2526
QY	1363	GCTAAGAAAGGACCAACATCAATGAAAGACTTAAGAAATCTCTGCTCTCTGACGCT	1422	QY	2443	CACACCTTTTCAGCAAGCAGCCCAAAAGGAGCAGCAACAGCTTTTGTGCTTAGC	2502
Db	1447	GCTAAGAAAGGACCAACATCAATGAAAGACTTAAGAAATCTCTGCTCTCTGACGCT	1506	Db	2527	CACACCTTTTCAGCAAGCAGCCCAAAAGGAGCAGCAACAGCTTTTGTGCTTAGC	2586
QY	1423	GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGAACATGAAGCAAAAGGT	1482	QY	2503	CCATGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTTAGATTTC	2562
Db	1507	GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGAACATGAAGCAAAAGGT	1566	Db	2587	CCATGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTTAGATTTC	2646
				QY	2563	AGCGGATGATGTCAGCGCTCTTCTGACAGGAGCCATGCCCCACTCTGCTCTGCTCTTG	2622


```
|||||
Db 2647 AGCAGATGATGTCAGCGCTCTTTCACAGCAGCCATGCCCCCATCTGCTCTGCCTCTTG 2706
QY 2623 TTACAGCCTCAAGTCTCAATAGTGTGAGAGCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2707 TTACAGGCTCAAGTGTCTCAATAGTGTGAGAGCCAGGAGCCACTGCAGATGCTCTCTC 2766
QY 2683 TTCAGGTCATCTAGCCCATCAAGCCTTCTGACGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2767 TTCAGGTCATCTAGCCCATCAAGCCTTCTGACGCCAGCAGTCTTGACAACTTATCTGG 2826
QY 2743 GAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAAACAGAGGCTGCTCCAGTTT 2802
Db 2827 GAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAAACAGAGGCTGCTCCAGTTT 2886
QY 2803 GGAGAAAAGAGGTTCCAGGAGTAGATTTAGCACTAACTCAATTCGTAAGAACTCTGG 2862
Db 2887 GGAGAAAAGAGGTTCCAGGAGTAGATTTAGCACTAACTCAATTCGTAAGAACTCTGG 2946
QY 2863 ACTTGAGCACCTAATGGATATATTTGACAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2947 ACTTGAGCACCTAATGGATATATTTGACAGAGAACAGATCACTTTGGATGTATTAGTTGA 3006
QY 2923 GATGGGCACAAAGGAGCTCAAGGAGATTGGAATCAATCTTATGGACATAGGCACAACT 2982
Db 3007 GATGGGCACAAAGGAGCTCAAGGAGATTGGAATCAATCTTATGGACATAGGCACAACT 3066
QY 2983 AATTAAGGAGTCGAGAGACTATCTCCGGACACAAAGCTTTAACCCTATTTAACTTT 3042
Db 3067 AATTAAGGAGTCGAGAGACTATCTCCGGACACAAAGCTTTAACCCTATTTAACTTT 3126
QY 3043 GAACACCTCTGTAGTGGAACTTCTTATAGATCTGTCTCTGATGATAAGAGTTTCA 3102
Db 3127 GAACACCTCTGTAGTGGAACTTCTTATAGATCTGTCTCTGATGATAAGAGTTTCA 3186
QY 3103 GTCTGTGAGGAAGAGATGCAAGTACAGTTTCAGAGTTCAGAGGTTTCTACAGAACTATG 3162
Db 3187 GTCTGTGAGGAAGAGATGCAAGTACAGTTTCAGAGTTCAGAGGTTTCTACAGAACTATG 3246
QY 3163 TGSAACTCTCAACAGATCAATATTTCTCAAGATTCAGAGTTCAGAGGTTTCTACAGAACTATG 3222
Db 3247 TGSAACTCTCAACAGATCAATATTTCTCAAGATTCAGAGTTCAGAGGTTTCTACAGAACTATG 3306
QY 3223 GGAAGATACACTCACCGGAGAAAAGAGTTTCTCAAGAAAACCAACCACTGCCAATGA 3282
Db 3307 GGAAGATACACTCACCGGAGAAAAGAGTTTCTCAAGAAAACCAACCACTGCCAATGA 3366
QY 3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTTATCCACAAAGCTTTGATGA 3342
Db 3367 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTTATCCACAAAGCTTTGATGA 3426
QY 3343 AAGGATGGGTACATAGTGGGTATGTTTGGAGCTGGCATTTATTTTGTGAAAACCTTTC 3402
Db 3427 AAGGATGGGTACATAGTGGGTATGTTTGGAGCTGGCATTTATTTTGTGAAAACCTTTC 3486
QY 3403 CAAAAGCAATCAATATGATATGTAATTCGAAATTCGAGGAGTACTGGGTCTCCAGTTTCACAAAGA 3462
Db 3487 CAAAAGCAATCAATATGATATGTAATTCGAAATTCGAGGAGTACTGGGTCTCCAGTTTCACAAAGA 3546
QY 3463 CAGATCTGTTTACATTTGTCACAGCAGCTGCTCTTTTGGCGGGTAACTTTGGGAAAGTC 3522
Db 3547 CAGATCTGTTTACATTTGTCACAGCAGCTGCTCTTTTGGCGGGTAACTTTGGGAAAGTC 3606
QY 3523 TTTTCTGCACTTTCAGTGCATCAAAATGGCAGATTCCTCCAGTGCATCACTAGTCAAC 3582
Db 3607 TTTTCTGCACTTTCAGTGCATCAAAATGGCAGATTCCTCCAGTGCATCACTAGTCAAC 3666
QY 3583 TGTAGGCCCACTGTAATGGCTTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACTA 3642
Db 3667 TGTAGGCCCACTGTAATGGCTTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACTA 3726
QY 3643 GCGTTATCTCGATATTTAATTTACTTACCAGATTTATGAGGCTGAAGGCTGATGGTGG 3702
|||||
```

```
Db 3727 GGCTTATCCTCGAGTATTTAATTACTTACCAGATTACGAGCCTGAAGCTATGTCGATGG 3786
QY 3703 ATAAATAGTTTATTTTAAAGAACTAAATTCCTCACTGAACCTAAATCATCAAGCAGCAGTGG 3762
Db 3787 ATAAATAGTTTATTTTAAAGAACTAAATTCCTCACTGAACCTAAATCATCAAGCAGCAGTGG 3846
QY 3763 CCTCTAGCTTTTACTCTCTTTGCTGAAAAAAA 3794
Db 3847 CCTCTAGCTTTTACTCCTTTGCTGAAAAAAA 3878

RESULT 6
AX062275
LOCUS AX062275
DEFINITION Sequence 134 from Patent WO0100849.
ACCESSION AX062275
VERSION AX062275.1 GI:12540176
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4992)
AUTHORS Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
TITLE Tankyrase2 materials and methods
JOURNAL Patent: WO 0100849-A 134 04 -JAN-2001;
ICOS CORPORATION (US)
FEATURES
Location/Qualifiers
source 1..4992
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 876..4376
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC25067.1"
/db_xref="GI:12540177"
/translation="MSRRRCAGGAGACASASAAAEEVPAARELFECRNGDVBVRKRLV
TPEKVNSDRTAGRKSTPLHFAAGFGKRDVVEYLLQANVQARDGGLIPLHNACSFQ
HAEVNLLRHGADPNARDNWTPLHEAAIKGKIDCVILVLLQHAEPTRIINTDGRTA
LDLADPSAKAVLTGEYKKDELLESARGNEEKMAALLTPLVNCHASDGRKSTPLILA
QTNPKVIYVOLLQHGADVHAKDGLVPLHNACSYGTHVEYELLVHKHACVNMDDLW
AGTPRHEAASKNRVEYSLLSYGDPTLLNCHNSAIDLAPTQPKRQICELLKRGAN
INKEKTEFTPLHVASEKAHNDVVEVVKHEAKNALNDLNGOTSILHRAAYCGHLOTCTR
LLLSYCGDNIISLOGFTALONGENVQQLQGGISLGNSEADRLLEAKGGLVPLHNA
KLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDGLVPLHNA
CSYGHVEAELLVLLQHGADVHAKDGLVPLHNAAGKYEICLLQHGADPTKKNRD
GNTPDLVDKGGTDIQLDLRGDAALLDAKGLARVKKLSPDNVNCRDQGRHSTP
LHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNA
TDKWAFTPLHEAAKGTOLCALLAHGADPTLKNQEGOTPLDLVADDSVALLTAAM
PPSLPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASISDLNLSGSELSVVS
SSGTGASLSLEKKEVPGVDFSTQFVRNLGLHLMDFERQITLDLVEMGHKELKE
IGNATVGHRLKLLIGVERLISGGQUNPYLTLLNTSGGTILIDLSPDPPKPEQSEEM
QSVTRHREHGGHAGGIFRNILIKQKCNKKLWERYTHRRKEVSENNHNHNRMLF
HGSFPFNATIIHKGFDERHAYIGMFGAGIYFAENSKNQYVYVGGGTGCPVHKDRS
CYICHRQLLPFRVTLLGKSLFQFSAMKMAHSPGHHSVTGRPSVNGLALEYVYRGEQ
AYPEYLITYQIMEPEGMDG"
BASE COUNT 1454 a 1064 c 1191 g 1283 t
ORIGIN
Query Match 89.2%; Score 3387.2; DB 6; Length 4992;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 403 AGGTTTTGGCGGGAAGACGCTAGTTGAATATTGCTTCAGAAATGGTCCAGTGTCCAAAGC 462
Db 1073 AGGTTTTGGCGGGAAGACGCTAGTTGAATATTGCTTCAGAAATGGTCCAAATGTCCAAAGC 1132
QY 463 ACGTGATCATGGGGCCCTTATCTCTTCATATGATGCTCTCTTTGCTCATGCTGAAGT 522
Db 1133 ACGTGATCATGGGGCCCTTATCTCTTCATATGATGCTCTCTTTGCTCATGCTGAAGT 1192
QY 523 AGTCAATCTCTCTTTTGGCAGATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATAC 582
```

Db 1193 AGTCAATCTCTTTTGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATAC 1252
QY 583 TCCTCTCCATGAAGCTCAATTAAGGAAAGATTGATGTTTGGATCTGCTGTGTACAGCA 642
Db 1253 TCCTCTCCATGAAGCTCAATTAAGGAAAGATTGATGTTTGGATCTGCTGTGTACAGCA 1312
QY 643 TGGAGCTGAGCCACCAATCCGAATACAGATGGAAGGACAGCATTTGGATTAGCAGATCC 702
Db 1313 TGGAGCTGAGCCACCAATCCGAATACAGATGGAAGGACAGCATTTGGATTAGCAGATCC 1372
QY 703 ATCTGCCAAGAGCTGCTTACTGCTGAATATAAGAAAGATGAACCTTTAGAAAGTGCCAG 762
Db 1373 ATCTGCCAAGAGCTGCTTACTGCTGAATATAAGAAAGATGAACCTTTAGAAAGTGCCAG 1432
QY 763 GAGTGGCAATGAAGAAAAATGATGGCTTACTCACACCATTAATTAATGCTCAACTGCCACGC 822
Db 1433 GAGTGGCAATGAAGAAAAATGATGGCTTACTCACACCATTAATTAATGCTCAACTGCCACGC 1492
QY 823 AAGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db 1493 AAGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1552
QY 883 TGTACAGCTCTTACTGCAACATGGAGCTGATGCTCCATGCTTAAAGATAAAGTGATCTGGT 942
Db 1553 TGTACAGCTCTTACTGCAACATGGAGCTGATGCTCCATGCTTAAAGATAAAGGTGATCTGGT 1612
QY 943 ACCATTACACAATGCCGTGTTCTTATGCTCATTTATGAAGTAACTGAACCTTTTGGTCAAGCA 1002
Db 1613 ACCATTACACAATGCCGTGTTCTTATGCTCATTTATGAAGTAACTGAACCTTTTGGTCAAGCA 1672
QY 1003 TGGTGCTGTGTAAGTCAATGACCTTGTGGCAATTCACCTCCTTCATGAGGACAGCTTC 1062
Db 1673 TGGTGCTGTGTAAGTCAATGACCTTGTGGCAATTCACCTCCTTCATGAGGACAGCTTC 1732
QY 1063 TAAGAACAGGGTTGAAGTATGTTCTTCTTCTTAAAGTTATGTTGGTGCAGACCCACACCTGCT 1122
Db 1733 TAAGAACAGGGTTGAAGTATGTTCTTCTTCTTAAAGTTATGTTGGTGCAGACCCACACCTGCT 1792
QY 1123 CAATTGTCACAATAAAGTGTCTATAGACTTGGCTCCACACACAGTATAAAGAAAGATT 1182
Db 1793 CAATTGTCACAATAAAGTGTCTATAGACTTGGCTCCACACACAGTATAAAGAAAGATT 1852
QY 1183 AGCATATGAATTTAAAGGCCACTGCTTGTGCAAGCTGCGACGAGAGCTGATGTTACTCG 1242
Db 1853 AGCATATGAATTTAAAGGCCACTGCTTGTGCAAGCTGCGACGAGAGCTGATGTTACTCG 1912
QY 1243 AATCAAAAAACATCTCTCTGGAATGGTGAATTTCAAGCATCCCTCAACACATGAAC 1302
Db 1913 AATCAAAAAACATCTCTCTGGAATGGTGAATTTCAAGCATCCCTCAACACATGAAC 1972
QY 1303 AGCATTGCATTTGCTGCTGATCTCCATATCCCAAAAGAAAGCAAAATATGTAACCTGTT 1362
Db 1973 AGCATTGCATTTGCTGCTGATCTCCATATCCCAAAAGAAAGCAAAATATGTAACCTGTT 2032
QY 1363 GCTAAGAAAGAGGACAAACATCAATGAAAGACTTAAGAAATTTTGAACCTCTGACAGCT 1422
Db 2033 GCTAAGAAAGAGGACAAACATCAATGAAAGACTTAAGAAATTTTGAACCTCTGACAGCT 2092
QY 1423 GGCATCTGAGAAAGCTCATTAATGATGTTGTTGAAGTAGTGGTGAAACATCAAGCAAGGT 1482
Db 2093 GGCATCTGAGAAAGCTCATTAATGATGTTGTTGAAGTAGTGGTGAAACATCAAGCAAGGT 2152
QY 1483 TAATGCTGTGATAAATTTGGTGAGACTTCTCTACAGAGCTGTCATATTTGGTGCATCT 1542
Db 2153 TAATGCTGTGATAAATTTGGTGAGACTTCTCTACAGAGCTGTCATATTTGGTGCATCT 2212
QY 1543 ACAAACCTGCCGCTTACTCTGAGCTATGGGTGATCTTAAACATTTATATCCCTTCAGGG 1602
Db 2213 ACAAACCTGCCGCTTACTCTGAGCTATGGGTGATCTTAAACATTTATATCCCTTCAGGG 2272
QY 1603 CTATTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCCTCCAGAGGGTATCTC 1662
Db 1662

Db 2273 CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCCTCCAGAGGGTATCTC 2332
QY 1663 ATTAGGTAATTCAGAGCGACAGACACAATTTGCTGGAAGCTGCCAAAGCTGGAGATGTCGA 1722
Db 2333 ATTAGGTAATTCAGAGCGACAGACACAATTTGCTGGAAGCTGCCAAAGCTGGAGATGTCGA 2392
QY 1723 AACTGTAAAAAAACTGTGTACTGTTACAGAGTGTCAACTGCGAGAGACATTGAAGGGCGTCA 1782
Db 2393 AACTGTAAAAAAACTGTGTACTGTTACAGAGTGTCAACTGCGAGAGACATTGAAGGGCGTCA 2452
QY 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGTGGGAATATCTGCT 1842
Db 2453 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGTGGGAATATCTGCT 2512
QY 1843 ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGGAGGCTTGTACCTTTTGCACAATGC 1902
Db 2513 ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGGAGGCTTGTACCTTTTGCACAATGC 2572
QY 1903 ATGTTCTTATGGACATTTATGAAGTTGCAGAACTTCTTTTAAACATGGAGCAGTAGTTAA 1962
Db 2573 ATGTTCTTATGGACATTTATGAAGTTGCAGAACTTCTTTTAAACATGGAGCAGTAGTTAA 2632
QY 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATGA 2022
Db 2633 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATGA 2692
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACCAAAAAAACAGGATGAAA 2082
Db 2693 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACCAAAAAAACAGGATGAAA 2752
QY 2083 TACTCCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2753 TACTCCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2812
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTTAGCCAGAGTGAAGAGTTGCTTC 2202
Db 2813 TGCAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTTAGCCAGAGTGAAGAGTTGCTTC 2872
QY 2203 TCCTGATTAATGTAATTTGCCCGCATACCCAAAGCAGACATTTCAACACCTTTTACATTTAGC 2262
Db 2873 TCCTGATTAATGTAATTTGCCCGCATACCCAAAGCAGACATTTCAACACCTTTTACATTTAGC 2932
QY 2263 AGCTGGTTATAATTTAGAACTTGCAGAGTATTTGTTTACAACACGAGCTGATGTTGAA 2322
Db 2933 AGCTGGTTATAATTTAGAACTTGCAGAGTATTTGTTTACAACACGAGCTGATGTTGAA 2992
QY 2323 TGCCCAAGACAAAGAGGAGCTTATTCCTTTTACATTAATGCGAGCATTTACCGGATGTAGA 2382
Db 2993 TGCCCAAGACAAAGAGGAGCTTATTCCTTTTACATTAATGCGAGCATTTACCGGATGTAGA 3052
QY 2383 TGTAGCAGCTTACTAATAAGTATAATGCTGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 3053 TGTAGCAGCTTACTAATAAGTATAATGCTGTGTCAATGCCACGGACAAATGGGCTTT 3112
QY 2443 CACACCTTTGCAGCAGAGCCCAAAAGGACGACACACAGCTTTTGTGCTTTGTTGCTAGC 2502
Db 3113 CACACCTTTGCAGCAGAGCCCAAAAGGACGACACACAGCTTTTGTGCTTTGTTGCTAGC 3172
QY 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTTACATTTAGTTTC 2562
Db 3173 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTTACATTTAGTTTC 3232
QY 2563 AGCGGATGATGTGACGGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTTGCCTCTTTG 2622
Db 3233 AGCAGATGATGTGACGGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTTGCCTCTTTG 3292
QY 2623 TTACAGCCTCAAGTGTCAATGTTGTGAGAAGCCAGGAGCCACTGCGATGCTCTC 2682
Db 3293 TTACAGCCTCAAGTGTCAATGTTGTGAGAAGCCAGGAGCCACTGCGATGCTCTC 3352
QY 2683 TTCAGTCTCATCTPAGCCCAAGCTTCTTGCAGCAGCAGCTTCTGCAGCAGCTTCTGCAGCAGCTTCTC 2742
Db 3353 TTCAGTCTCATCTPAGCCCAAGCTTCTTGCAGCAGCAGCTTCTGCAGCAGCTTCTGCAGCAGCTTCTC 3412

Qy	2743	GAGTTTTTCAGAACTGCTCTTCAGTAGTTAGTTCAGTGGAAACAGAGGGTGCTTCCAGTTT	2803
Db	3413	GAGTTTTTCAGAACTGCTCTTCAGTAGTTAGTTCAGTGGAAACAGAGGGTGCTTCCAGTTT	3472
Qy	2803	GGAAAAAGGAGTTCCAGGAGTAGTTTTAGCATCAACTCAATTCGTAAAGGAATCTTGG	2862
Db	3473	GGAAAAAGGAGTTCCAGGAGTAGTTTTAGCATCAACTCAATTCGTAAAGGAATCTTGG	3532
Qy	2863	ACTTGAGCACCTAATGGATATATTTGAGAGAAACAGATCACTTTGGATGATATTAGTTGA	2922
Db	3533	ACTTGAGCACCTAATGGATATATTTGAGAGAAACAGATCACTTTGGATGATATTAGTTGA	3592
Qy	2923	GATGGGCGACAAGAGCTGAAGAGATTGGAAATCAATGCTTTATGGACATAGGCACAAACT	2982
Db	3593	GATGGGCGACAAGAGCTGAAGAGATTGGAAATCAATGCTTTATGGACATAGGCACAAACT	3652
Qy	2983	AAITAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGCTTAACCCATATTATTAACTTT	3042
Db	3653	AAITAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGCTTAACCCATATTATTAACTTT	3712
Qy	3043	GAACACCTCTGCTAGTGGAAACAATTCCTTATAGATCTGCTCCTCATGATATAAGAGTTTCA	3102
Db	3713	GAACACCTCTGCTAGTGGAAACAATTCCTTATAGATCTGCTCCTCATGATATAAGAGTTTCA	3772
Qy	3103	GTCGTGGAGGAAGAGATGCAAAAGTACAGTTTCGAGAGCACAGAGATGGAGTCAATGCAGG	3162
Db	3773	GTCGTGGAGGAAGAGATGCAAAAGTACAGTTTCGAGAGCACAGAGATGGAGTCAATGCAGG	3832
Qy	3163	TGGAATCTTCAACAGATACAAATATTTCTCAAGATTTCAGAAAGTTGTATACAAAGAACTATG	3222
Db	3833	TGGAATCTTCAACAGATACAAATATTTCTCAAGATTTCAGAAAGTTGTATACAAAGAACTATG	3892
Qy	3223	GGAAGATACACTCACCGGAGAAAAGAGTTCTTCTGAAGAAAACACCAACCATGCGCAATGA	3282
Db	3893	GGAAGATACACTCACCGGAGAAAAGAGTTCTTCTGAAGAAAACACCAACCATGCGCAATGA	3952
Qy	3283	ACGAATGCTATTTTCATGGGTCTCCTTTTCTGAATGCAATTTATCCAAAGGCTTTGATGA	3342
Db	3953	ACGAATGCTATTTTCATGGGTCTCCTTTTCTGAATGCAATTTATCCAAAGGCTTTGATGA	4012
Qy	3343	AAGCATCGGTACATAGTGGTATGTTTCGAGCTGGCATTTATTTTGGTGAAGAACTCTTC	3402
Db	4013	AAGCATCGGTACATAGTGGTATGTTTCGAGCTGGCATTTATTTTGGTGAAGAACTCTTC	4072
Qy	3403	CAAAAGCAATCAATATGATATGGAATTCGGAGAGGTACTGGGTGCCAGTTTCAAAAGA	3462
Db	4073	CAAAAGCAATCAATATGATATGGAATTCGGAGAGGTACTGGGTGCCAGTTTCAAAAGA	4132
Qy	3463	CAGATCTTGTATTTTCAGAGGAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTC	3522
Db	4133	CAGATCTTGTATTTTCAGAGGAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTC	4192
Qy	3523	TTTCTCGAGTTTCAGTGCATGAAATGCAATTCCTCTCCAGTGCATCACTCAGTCAAC	3582
Db	4193	TTTCTCGAGTTTCAGTGCATGAAATGCAATTCCTCTCCAGTGCATCACTCAGTCAAC	4252
Qy	3583	TGGTAGGCCAGTGAATGGCCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAACA	3642
Db	4253	TGGTAGGCCAGTGAATGGCCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAACA	4312
Qy	3643	GCCTATTCCTCAGTATTTAATTTACTTACAGATTTATGAGGCCCTGAAGGTATGGTCGATGG	3702
Db	4313	GCCTATTCCTCAGTATTTAATTTACTTACAGATTTATGAGGCCCTGAAGGTATGGTCGATGG	4372
Qy	3703	ATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTAAAATCATCAAGCAGCAGTGG	3762
Db	4373	ATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTAAAATCATCAAGCAGCAGTGG	4432
Qy	3763	CCTCTACGTTTTTACTTCCTTTGCTGAAAAAAA	3794
Db	4433	CCTCTACGTTTTTACTTCCTTTGCTGAAAAAAA	4464

RESULT 7	AX062273	Sequence	132 from Patent WO0100849.	5002 bp	DNA	linear	PAT 24-JAN-2001		
LOCUS	AX062273	DEFINITION	Sequence 132 from Patent WO0100849.						
ACCESSION	AX062273	KEYWORDS	AX062273.1	GI:12540174					
VERSION		SOURCE	human.						
ORGANISM			Homo sapiens						
REFERENCE			Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
AUTHORS			Christenson, E., Demaggio, A.J., Goldman, P.S. and Mcelligott, D.L.						
TITLE			Tankyrase2 materials and methods						
JOURNAL			Patent: WO 0100849-A 132 04-JAN-2001;						
FEATURES			ICOS CORPORATION (US)						
source			Location/Qualifiers						
			1..5002						
			/organism="Homo sapiens"						
			/db_xref="taxon:9606"						
CDS			229..4386						
			/note="unnamed protein product"						
			/codon_start=1						
			/protein_id="C25066.1"						
			/db_xref="GI:12540175"						
			/translation="MMKLKTRTISTVRLPKECARSEQNQGLFFSVVYRKFSLESFPG LKTNTKNTICRIVRQONPQRWRDEAPSPAAEASRLTPTNPKRRLASRLNLPSC PATAHARDYDTCAPGARSQWELAGALPASAAASQDDPGDFALPPPPRSGRAGSP ARGARGRGHTAPDPVTAGSOAARLASSPGGLALLLAGPGLLRLLALLLAVAAAR IMSGRRACAGGAACAASAAEAPPAARELFACRNGDVERVKRLVTPKVNRSRTAGR KSTPLHPAAGFGPKDVEYLWLCANVOARDGGGLPLHNACSEFGHAENVLLRHGA DPNARDNWTPLHEAAAKGKIDVICVLQHGAEPTIRTDGRTALDLADPSKAVLT QGYKDELLSASRSGNEEKMMALLPLNVNCHASDGRKSTPLHLAAGIRVKIKVLQLL QHQADYHAKDQGDPLVPLHNACSYGHEVTELLYKHGACVNAMDWQFTPLHEAASKN VEVCSLLLSYGADPTLHNCHKSAIDLAFTPOLKERLAYEFKGSLLQAAAREADVTRI KKHLSLEMVNFKHPOTLHNCHKSAIDLAFTPOLKERLAYEFKGSLLQAAAREADVTRI VASEKANDVENVVVKHEAKVNALDNGOTSLHRAAYCGHLOTICRLLSYGCDPNIIIS LQGTALQGNFVNVQQLQBGISLGNSEADQRLLEAKKAGDVETVKKLTQVQNCRD TEGRSTPLHPAAGFGPKDVEYLWLCANVOARDGGGLPLHNACSEFGHAENVLLRHGA KHGAVNVADLWFTPLHEAAAKGYEICKLLQHGADPTKKNRDNTPDLIVKDGDT DIODPLRGDAALLDAARKGLCLARVKLSSPDNVNCRDTPGRHSTPLHLAAGIRNVLEVA EYLLQHCADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFPLHEAA OKGTLOCALLLAHAGDPTLKNQGGTPLDVSADDDVALLTAAMPSPALPCYKQPV LNGVRSPTADALSSGPPSPSLSAASLDLNSGFSELSVSVSSGTGEGASLLEKK EVPVDISITQFVNLGUEHMDIFEREQITDVLVEMGHKEKEIGINAYGHRHGLI KGVRLITSGQGLNPLTINTSGGTLIDLDSPDKFEQFVEEMQSTVEYHRRHGLI GFIENYILKIQVCNKKLWENSYTHRKVESBENHNHANERMLFHGSPFVNAIIHK FDERHAYIGFMGAGIYFAENSSKNQYVYIGGTCGCPVKDKRSCVICHRLQLFCRV TLCKSFLOFSAMKMAHSPPGHHVSTGPRSVNGLALAEVYIRGEQAPVEYLITQIMLR PEGWVDG"						
BASE COUNT			1456 a	1065 c	1196 g	1285 t			
ORIGIN									
			Quality Match	89.24;	Score 3387.2;	DB 6;	Length 5002;		
			Best_Local Similarity	99.94;	Pred. No.:				
			Matches 3389;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;		
Qy	403	AGGTTTGGGCGGAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAGATGTCACAAGC	462						
Db	1083	AGGTTTGGGCGGAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAGAAATGTCACAAGC	1142						
Qy	463	AGCTGATGATGGGGCCCTTATTCCTTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT	522						
Db	1143	AGCTGATGATGGGGCCCTTATTCCTTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT	1202						
Qy	523	AGTCAATCPCCTTTTGGCACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC	582						
Db	1203	AGTCAATCPCCTTTTGGCACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC	1262						
Qy	583	TCCTCTCCATGAAGTCGCAATTTAAAGGAAAGATTGATGCTTTTGCAATGTGCTGTGTACAGCA	642						
Db	1263	TCCTCTCCATGAAGTCGCAATTTAAAGGAAAGATTGATGCTTTTGCAATGTGCTGTGTACAGCA	1322						

QY 643 TGGAGCTGAGCCAAACATCCGAAATACAGATGAAGACAGCATTTGATTTAGCAGATCC 702
|||||
Db 1323 TGGAGCTGAGCCAAACATCCGAAATACAGATGAAGACAGCATTTGATTTAGCAGATCC 1382
QY 703 ATCTGCCAAGCAGTCTGTTACTGCTGAATATGAAGAAAGATGAACCTTTAGAAAGTGCAG 762
|||||
Db 1383 ATCTGCCAAGCAGTCTGTTACTGCTGAATATGAAGAAAGATGAACCTTTAGAAAGTGCAG 1442
QY 763 GAGTGGCAATGAAGAAAAATGATGCTCTACTCACACCATTAATGTCAACTGCCACGC 822
|||||
Db 1443 GAGTGGCAATGAAGAAAAATGATGCTCTACTCACACCATTAATGTCAACTGCCACGC 1502
QY 823 AAGTGTGCGAAGTCAACTCCATTAATTTGCGACAGAGATATAACAGAGTAAAGAT 882
|||||
Db 1503 AAGTGTGCGAAGTCAACTCCATTAATTTGCGACAGAGATATAACAGAGTAAAGAT 1562
QY 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGCTCCATGCTTAAGATTAAGTGTCTGGT 942
|||||
Db 1563 TGTACAGCTGTTACTGCAACATGGAGCTGATGCTCCATGCTTAAGATTAAGTGTCTGGT 1622
QY 943 ACCATTACAAATGCTGTTTATGGTCATTAATGAAGTAACTGAACCTTTGGTCAAGCA 1002
|||||
Db 1623 ACCATTACAAATGCTGTTTATGGTCATTAATGAAGTAACTGAACCTTTGGTCAAGCA 1682
QY 1003 TGGTGCCTGTAAATGCAATGGACTTGGGCAATTCACCTCTTCATGAGGCAAGCTTC 1062
|||||
Db 1683 TGGTGCCTGTAAATGCAATGGACTTGGGCAATTCACCTCTTCATGAGGCAAGCTTC 1742
QY 1063 TAAGAACAGGTTGAAGTATGTTCTCTCTCTTAAGTTATGTTGCGAGACCCCAACACTGCT 1122
|||||
Db 1743 TAAGAACAGGTTGAAGTATGTTCTCTCTCTTAAGTTATGTTGCGAGACCCCAACACTGCT 1802
QY 1123 CAATTGTCAATAAAGTGCTATAGACTTGGCTCCCAACACAGTAAAGAAAGATT 1182
Db 1803 CAATTGTCAATAAAGTGCTATAGACTTGGCTCCCAACACAGTAAAGAAAGATT 1862
QY 1183 AGCATATGAATTAAGGCCACTCTGCTGCAAGCTCCAGAGAGCTGATGTTACTCG 1242
Db 1863 AGCATATGAATTAAGGCCACTCTGCTGCAAGCTCCAGAGAGCTGATGTTACTCG 1922
QY 1243 AATCAAAAACATCTCTCTGGAATGGTGAATTTCAAGCATCTCAACACATGAAC 1302
Db 1923 AATCAAAAACATCTCTCTGGAATGGTGAATTTCAAGCATCTCAACACATGAAC 1982
QY 1303 AGCATATGATTTGCTGCTGATCTCCATATCCCAAGAAAGCAATATGTGAACCTGT 1362
Db 1983 AGCATATGATTTGCTGCTGATCTCCATATCCCAAGAAAGCAATATGTGAACCTGT 2042
QY 1363 GCTAAGAAAGGAGCAACATCAATGAAGACTAAAGAAATTTGACTCTCTCTGCACT 1422
Db 2043 GCTAAGAAAGGAGCAACATCAATGAAGACTAAAGAAATTTGACTCTCTCTGCACT 2102
QY 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTGTGTTGAAGTGAAGCAATGAAGCAAGGT 1482
Db 2103 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTGTGTTGAAGTGAAGCAATGAAGCAAGGT 2162
QY 1483 TAATGCTCTGATATCTGCTGCTGACTTCTACACAGAGCTGCATATGTTGCTCATCT 1542
Db 2163 TAATGCTCTGATATCTGCTGCTGACTTCTACACAGAGCTGCATATGTTGCTCATCT 2222
QY 1543 ACAAACTTCCGCTTACTCTCTGAGCTATGGGTGTGATCCCTAACATTAATCCCTTCAGG 1602
Db 2223 ACAAACTTCCGCTTACTCTCTGAGCTATGGGTGTGATCCCTAACATTAATCCCTTCAGG 2282
QY 1603 CTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCACTCTCCCAAGAGGATATCTC 1662
Db 2283 CTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCACTCTCCCAAGAGGATATCTC 2342
QY 1663 ATTAGGTAATTCAGGCGACAGACAAATGCTGCGAAGCTGCAAGGCTGGAGATGCGA 1722
Db 2343 ATTAGGTAATTCAGGCGACAGACAAATGCTGCGAAGCTGCAAGGCTGGAGATGCGA 2402

QY 1723 AACTGTAAAAAACTGTGTACTGTTTACAGAGTGTCACTGCAGAGACATTTGAAGGGCGTCA 1782
|||||
Db 2403 AACTGTAAAAAACTGTGTACTGTTTACAGAGTGTCACTGCAGAGACATTTGAAGGGCGTCA 2462
QY 1783 GTCATACCACTTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCT 1842
Db 2463 GTCATACCACTTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCT 2522
QY 1843 ACAGCATGGAGCTGATGCTGCTAAAGATAAAGAGGCTTGTACCTTTGCACATGC 1902
Db 2523 ACAGCATGGAGCTGATGCTGCTAAAGATAAAGAGGCTTGTACCTTTGCACATGC 2582
QY 1903 ATGCTCTTATGCGACATTAATGAAGTTGCGAAGCTTCTGTTTAAACATGGAGCAGTGTGTTAA 1962
Db 2583 ATGCTCTTATGCGACATTAATGAAGTTGCGAAGCTTCTGTTTAAACATGGAGCAGTGTGTTAA 2642
QY 1963 TGTAGCTGATTTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAAGAAATATGA 2022
Db 2643 TGTAGCTGATTTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAAGAAATATGA 2702
QY 2023 AATTGCAAACTTCTGCTCCAGCATGCTGCGAGACCTTACCAAAAACAGGATGGA 2082
Db 2703 AATTGCAAACTTCTGCTCCAGCATGCTGCGAGACCTTACCAAAAACAGGATGGA 2762
QY 2083 TACTCTCTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 2142
Db 2763 TACTCTCTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 2822
QY 2143 TGCAGCTTTCCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAGTGTCTTC 2202
Db 2823 TGCAGCTTTCCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAGTGTCTTC 2882
QY 2203 TCTCTGATTAATGTAATTTGCCGCGATACCCAAAGCAGACATTTCAACACCTTTACATTTAGC 2262
Db 2883 TCTCTGATTAATGTAATTTGCCGCGATACCCAAAGCAGACATTTCAACACCTTTACATTTAGC 2942
QY 2263 AGCTGGTTTAAATTAATTTAGAAGTTCAGAGTATTTGTTTACAAACAGGAGCTGATGTGAA 2322
Db 2943 AGCTGGTTTAAATTAATTTAGAAGTTCAGAGTATTTGTTTACAAACAGGAGCTGATGTGAA 3002
QY 2323 TGCCCAAGCAAAAGGAGGACTTATCTTTACATAATGAGCATCTTACGGGCATGTAGA 2382
Db 3003 TGCCCAAGCAAAAGGAGGACTTATCTTTACATAATGAGCATCTTACGGGCATGTAGA 3062
QY 2383 TGTAGCAGCTCTACTAATAAAGTATTAATGCTGTCATGTCATGTCACGACACAAATGGGCTTT 2442
Db 3063 TGTAGCAGCTCTACTAATAAAGTATTAATGCTGTCATGTCATGTCACGACACAAATGGGCTTT 3122
QY 2443 CACACCTTTGCGAGAGCAGCCCAAAAGGAGCAACACAGCTTGTGCTGTAGC 2502
Db 3123 CACACCTTTGCGAGAGCAGCCCAAAAGGAGCAACACAGCTTGTGCTGTAGC 3182
QY 2503 CCATGAGCTGAGCCGACTCTTAAATAATCAGGAAGGAGCAACACCTTTAGATTTAGTTTC 2562
Db 3183 CCATGAGCTGAGCCGACTCTTAAATAATCAGGAAGGAGCAACACCTTTAGATTTAGTTTC 3242
QY 2563 AGCGGATGATGTGAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTCCCTCTTG 2622
Db 3243 AGCAGATGATGTGAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTCCCTCTTG 3302
QY 2623 TTACACACCTCAAGTCTCAATGGTGTGAGAGCCAGGAGCCACTGCGATGCTCTCTC 2682
Db 3303 TTACACACCTCAAGTCTCAATGGTGTGAGAGCCAGGAGCCACTGCGATGCTCTCTC 3362
QY 2683 TTTAGCTTCCATCTAGCCCATCAAGCTTTCTGCGAGCAGCAGCTTTGACAACTTATCTGG 2742
Db 3363 TTTAGCTTCCATCTAGCCCATCAAGCTTTCTGCGAGCAGCAGCTTTGACAACTTATCTGG 3422
QY 2743 GAGTGTTCAGAACTGTCTTCAAGTATGTTAGTTCAAGTGAACAGAGGCTGCTTCCAGTTT 2802
Db 3423 GAGTGTTCAGAACTGTCTTCAAGTATGTTAGTTCAAGTGAACAGAGGCTGCTTCCAGTTT 3482
QY 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCAATACTCAATTCGTAAGGAATCTTGG 2862

Db 3483 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATACTCAATTCGTAAGGAATCTTGG 3542
QY 2863 ACTTGAGCACCTTAATGGATATATTTGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 3543 ACTTGAGCACCTTAATGGATATATTTGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3602
QY 2923 GATGGGGCACAGGAGCTGAGGAGATTGGAATCAATGCTTTATGGACATAGGCACAACT 2982
Db 3603 GATGGGGCACAGGAGCTGAGGAGATTGGAATCAATGCTTTATGGACATAGGCACAACT 3662
QY 2983 AATTAAAGGAGTCGAGACATTTATCTCCGGACACAGGCTTTAACCATAATTTAACTTT 3042
Db 3663 AATTAAAGGAGTCGAGACATTTATCTCCGGACACAGGCTTTAACCATAATTTAACTTT 3722
QY 3043 GAACACCTCTGGTAGTGAACAATTTCTTATAGATCTGCTCCTCATGATGAAGAGTTTCA 3102
Db 3723 GAACACCTCTGGTAGTGAACAATTTCTTATAGATCTGCTCCTCATGATGAAGAGTTTCA 3782
QY 3103 GTCTGTGAGCAAGAGATGCAAACTACAGTCGAGACACAGAGATGGAGTTCATGCAGG 3162
Db 3783 GTCTGTGAGCAAGAGATGCAAACTACAGTCGAGACACAGAGATGGAGTTCATGCAGG 3842
QY 3163 TGGAACTCTTCAACAGATACATATTTCAAGATTCAAGAGTTTGTAAACAGAACTATG 3222
Db 3843 TGGAACTCTTCAACAGATACATATTTCAAGATTCAAGAGTTTGTAAACAGAACTATG 3902
QY 3223 GGAAGATACACTCACCGGAGAAAGATTTCTGAAGAAACACCAACCATGCCAATGA 3282
Db 3903 GGAAGATACACTCACCGGAGAAAGATTTCTGAAGAAACACCAACCATGCCAATGA 3962
QY 3283 ACGAATGCTATTTATGCGGCTCTCTTTGGAATGCAATATATCCAAAGGCTTTGATGA 3342
Db 3963 ACGAATGCTATTTATGCGGCTCTCTTTGGAATGCAATATATCCAAAGGCTTTGATGA 4022
QY 3343 AAGCATGCTACATAGTGGTATGTTGGAGCTGGCATTTATTTTGGTGAACACTCTTC 3402
Db 4023 AAGCATGCTACATAGTGGTATGTTGGAGCTGGCATTTATTTTGGTGAACACTCTTC 4082
QY 3403 CAAAGCAATCAATATCTATGGAATGGAGAGTACTGGGTGTCAGTTCCAGTTCCAAAGA 3462
Db 4083 CAAAGCAATCAATATCTATGGAATGGAGAGTACTGGGTGTCAGTTCCAGTTCCAAAGA 4142
QY 3463 CAGATCTTGTTACATTTGCCACAGGAGCTGCTCTTTTCCGGGTAACTTCCGGAAGTC 3522
Db 4143 CAGATCTTGTTACATTTGCCACAGGAGCTGCTCTTTTCCGGGTAACTTCCGGAAGTC 4202
QY 3523 TTTCTCTGAGTTCAGTGCATGAATGCAATGCGACATTTCTCTCCAGGTCACTCAGTCAC 3582
Db 4203 TTTCTCTGAGTTCAGTGCATGAATGCAATGCGACATTTCTCTCCAGGTCACTCAGTCAC 4262
QY 3583 TGGTAGGCCAGTGTAAATGGCTAGCATTAGCTGAATATGAGGCTGAAGTATGTCGATGG 3702
Db 4263 TGGTAGGCCAGTGTAAATGGCTAGCATTAGCTGAATATGAGGCTGAAGTATGTCGATGG 4322
QY 3643 GCGTATCTGAGTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3702
Db 4323 GCGTATCTGAGTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4382
QY 3703 ATAAATAGTATTTTAAAGAACTAATTCACCTGAACCTTAAATCAATCAATCAATCAAT 3762
Db 4383 ATAAATAGTATTTTAAAGAACTAATTCACCTGAACCTTAAATCAATCAATCAATCAAT 4442
QY 3763 CCTCTACGTTTACCTCTTCTGCTGAAAAAAA 3794
Db 4443 CCTCTACGTTTACCTCTTCTGCTGAAAAAAA 4474

RESULT 8

AF264912 5810 bp mRNA linear PRI 14-MAY-2001
LOCUS Homo sapiens tankyrase-like protein (TNKL) mRNA, complete cds.
DEFINITION AF264912
ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

AF264912.1 GI:12005975

Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 5810)

Kulmov, A.N., Kuprash, D.V., Petrov, V.N., Vdovichenko, K.K.,

Scanlan, M.J., Jongeneel, C.V., Lagarkova, M.A., and Nedospasov, S.A.

Cloning and characterization of TNKL, a member of tankyrase gene

family

Genes Immun. 2 (1), 52-55 (2001)

21190090

11294570

2 (bases 1 to 5810)

Kulmov, A.N., Kuprash, D.V., Petrov, V.N., Vdovichenko, K.K.,

Kashikarova, U.A., Boitchenko, V.E., Nedospasov, S.A. and

Lagarkova, M.A.

Direct Submission

Submitted (08-MAY-2000) Laboratory of Molecular Immunology, A.N.

Belozersky Institute of Physico-Chemical Biology, Moscow State

University, Moscow 119899, Russia

Location/Qualifiers

1..5810

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="10"

/map="10q22-10q24"

/tissue_type="breast carcinoma"

/note="derived from clones MO-BC-203 and I203"

<1..5810

/gene="TNKL"

/note="a new member of tankyrase gene family; current

UniGene cluster Hs.168491"

1..3501

/gene="TNKL"

/note="a new member of tankyrase family"

/codon_start=1

/product="tankyrase-like protein"

/protein_id="AAG44694.1"

/db_xref="GI:12005976"

/translation="MSGRRCAAGGACACASAAAEVPAARELFECARNGDVERKRLV

TPEKNSRDTAGKSTPLHFAAGFGRKDVVEYLONGANVQARDGGLPLHNCASF

HAENVNLLRHGADPNARDNNVTPLEAAIKKIDVICVILQHGAEPTIRNTDGRTA

LDLADPSAKAVLTGEYKKDELSARGNEEKMMALLPLNVNCHASDGRKSTPLHLA

AGYRVKIVQLLQHGADVHAKDGLVPLHNCASGHHYEVELLYKHGACVNMADLW

QFTPLHEADNRKVEVCSLLSYGADPTLLNCHNSAIDLAFTPOLKRLAYEFKQHS

LLOAAREADVTRIKHLSEMVNFQHTETALHCAASAPYPRKQICPLLRKAN

INFTKEFTPLHVASEKAHNDVVEVVKHEAKVNDLNGOTSLHRAAYCGHQTGR

LLSYGCDPNIISLQGTALQMGNEVQQLQSGISLGNSEADRLLEAAKAGDVEV

KKLCTVQSNCRDLEGROSTPLHFAAGYNRVSVVEILLQHGADVHAKDGLVPLHNA

CSYGHVEALLVKGAVNVADLWRTPLHEAAAGKYEICKLLQHGADPTKKNRD

GNTPDLVDKGDITDIDLRGDAALDAKAGCLARVKLSLSPDNVNCRDGTGKRLN

LHLAAGTNLVEAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNA

PLWAFTPLHEAAKGTQICALLAHGADPTLLNCHNSAIDLAFTPOLKRLAYEFKQHS

PPSALPCYKQVINGVRSGATADALSSGSPSSLSAASIDNLNLSGFSSELSSVVS

SSGTEGASSLEKKEVPGVDFSTIQVFNGLGLEHLMDFPERQITLIDLVMGHKELKE

IGINAYGHRHKLKIGVERLISGQGLNPLILTNTSGSTLIDLSPDDKEPQSVBEEM

QSVREHRRDGGHAGGIFNRYNLIKQKCNKKLWERYTHRRKEVSEHNHNERMLF

HSGFPVNAIHRGFDERHAYIGFMFAGIYFAENSNSKNOYVIGGGTGCPCVHKDRS

CICHKQLLFCRVTLTKSFLQFSAMKMAHSPPHSHSVTGRPSVNGLALAEYVYIRGEQ

AYPEYLITYQIMRPEGVDC"

complement(1132..1277)

/note="SHGC-53837"

complement(3665..3975)

/note="WI-16054"

complement(5544..5777)

/note="WI-8339"

BASE COUNT 1736 a 1094 c 1231 g 1749 t

ORIGIN

Query Match

Best Local Similarity

89.2%; Score 3387.2; DB 9; Length 5810;

99.9%; Pred. No. 0;

Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
Qy	403	AGGTTTGGCGGAAAGACGTAGTTCGAATATTTGCTTCAGAAATGGTGCAGAGTTCGCAAGC	462
Db	198	AGGTTTGGCGGAAAGACGTAGTTCGAATATTTGCTTCAGAAATGGTGCAGAAATGCTCCAAGC	257
Qy	463	ACGTGATGATGGGGGCTTATTCCTCTTCATATGATGCTCTTTTGGTCATGCTGGAAGT	522
Db	258	ACGTGATGATGGGGCTTATTCCTCTTCATATGATGCTCTTTTGGTCATGCTGGAAGT	317
Qy	523	AGTCAATCTCTTTTGGCAGATGGTCAGACCCCAATGCTCGAGATAAATTTGGAATATATAC	582
Db	318	AGTCAATCTCTTTTGGCAGATGGTCAGACCCCAATGCTCGAGATAAATTTGGAATATATAC	377
Qy	583	TCCFTCCATGAAGCTGCAATTAAGAGAAAGATGTATGTTTGCAATTTGCTGCTTTACAGCA	642
Db	378	TCCTCTCCATGAAGCTGCAATTAAGAGAAAGATGTATGTTTGCAATTTGCTGCTTTACAGCA	437
Qy	643	TGGAGCTGAGCCACCAATCCGAATATACAGATGGAAGCAGACAGCATTTGATTTACAGATCC	702
Db	438	TGGAGCTGAGCCACCAATCCGAATATACAGATGGAAGCAGACAGCATTTGATTTACAGATCC	497
Qy	703	ATCTGCCAAAGCAGTCTACTGCTGTAATATAAGAAAGATGAACCTCTTAGAAAAGTGCCAG	762
Db	498	ATCTGCCAAAGCAGTCTACTGCTGTAATATAAGAAAGATGAACCTCTTAGAAAAGTGCCAG	557
Qy	763	GAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTAAATGCTCAACTGCCACGC	822
Db	558	GAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTAAATGCTCAACTGCCACGC	617
Qy	823	AGTGCATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGATFAAAGAT	882
Db	618	AGTGCATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGATFAAAGAT	677
Qy	883	TGTACAGCTGTTACTGCAACATGGAGCTGATGCCATGCTTAAGATAAAGGTGATCTGGT	942
Db	678	TGTACAGCTGTTACTGCAACATGGAGCTGATGCCATGCTTAAGATAAAGGTGATCTGGT	737
Qy	943	ACCAATTACACAATGCTGTCTTATGTTGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA	1002
Db	738	ACCAATTACACAATGCTGTCTTATGTTGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA	797
Qy	1003	TGGTGGCTGTGTAATGCAATGGACTTGTGGCAATTTCACTCCTCTCATGAGGCAGCTTC	1062
Db	798	TGGTGGCTGTGTAATGCAATGGACTTGTGGCAATTTCACTCCTCTCATGAGGCAGCTTC	857
Qy	1063	TAAAGACAGGGTGAAGTATGTCCTCTCTTAAGTTATGGTCGAGACCCCAACACTGCT	1122
Db	858	TAAAGACAGGGTGAAGTATGTCCTCTCTTAAGTTATGGTCGAGACCCCAACACTGCT	917
Qy	1123	CAATTGTCACAATAAAGTGCTATAGACTTGGCTCCCACACACAGTTAAAGAAAGATT	1182
Db	918	CAATTGTCACAATAAAGTGCTATAGACTTGGCTCCCACACACAGTTAAAGAAAGATT	977
Qy	1183	AGCATATGAATTTAAGGCCACTCGTTGCTGCAAGCTGCACAGAAAGCTGATTTACTCG	1242
Db	978	AGCATATGAATTTAAGGCCACTCGTTGCTGCAAGCTGCACAGAAAGCTGATTTACTCG	1037
Qy	1243	AATCAAAAAACATCTCTCTGSAATTTGTAATTTCAAGCATCCTCAACACATGAAC	1302
Db	1038	AATCAAAAAACATCTCTCTGSAATTTGTAATTTCAAGCATCCTCAACACATGAAC	1097
Qy	1303	AGCATATGATTTGCTGTCATCTCCATATCCCAAAAAGAAAGCAAAATGTGAACGTGT	1362
Db	1098	AGCATATGATTTGCTGTCATCTCCATATCCCAAAAAGAAAGCAAAATGTGAACGTGT	1157
Qy	1363	GCTAAGAAAAAGGACAAACATCAATGAAAGACTAAGAAATTTCTTGACTCCTCTGCACGT	1422
Db	1158	GCTAAGAAAAAGGACAAACATCAATGAAAGACTAAGAAATTTCTTGACTCCTCTGCACGT	1217
Qy	1423	GGCATCTGAGAAAGCTCATTAATGATGTTGTTGAAGTAGTGGTGAACACATCAAGCAAGCT	1482
Db	1218	GGCATCTGAGAAAGCTCATTAATGATGTTGTTGAAGTAGTGGTGAACACATCAAGCAAGGT	1277

QY 2683 TTCAGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTTATCTGG 2742
DB 2857 TTCAGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTTATCTGG 2916
QY 2743 GAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGGTGTCTCCAGTTT 2802
DB 2917 GAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGGTGTCTCCAGTTT 2976
QY 2803 GGAGAAAAGAGGTTCCAGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
DB 2977 GGAGAAAAGAGGTTCCAGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 3036
QY 2863 ACTTCAGCACCTAATGGATATATTGGAGAGAACAGATCATCTTTGGATGTATTAGTTGA 2922
DB 3037 ACTTCAGCACCTAATGGATATATTGGAGAGAACAGATCATCTTTGGATGTATTAGTTGA 3096
QY 2923 GATGGGCACAAAGAGCTGAAGAGATTTGAATCAATGCTTTATGACATAGGCACAACT 2982
DB 3097 GATGGGCACAAAGAGCTGAAGAGATTTGAATCAATGCTTTATGACATAGGCACAACT 3156
QY 2983 AATTAAAGGAGTCGAGACATTTATCTCCGGACAACAAGGCTTTAACCATATTTAACTTT 3042
DB 3157 AATTAAAGGAGTCGAGACATTTATCTCCGGACAACAAGGCTTTAACCATATTTAACTTT 3216
QY 3043 GAACACCTCTGTGTAGTGAACAATTTATAGATCTGTCTCTGATGATATAAGAGTTTCA 3102
DB 3217 GAACACCTCTGTGTAGTGAACAATTTATAGATCTGTCTCTGATGATATAAGAGTTTCA 3276
QY 3103 GTCTGTGAGGAAGATGCAAACTACAGTTTCGAGAGCACAGATGGAGTCTATGAGG 3162
DB 3277 GTCTGTGAGGAAGATGCAAACTACAGTTTCGAGAGCACAGATGGAGTCTATGAGG 3336
QY 3163 TGGAAATCTTCACAGATACATATTTCTCAAGATTCAGAGGTTTTGTACACAGAATATG 3222
DB 3337 TGGAAATCTTCACAGATACATATTTCTCAAGATTCAGAGGTTTTGTACACAGAATATG 3396
QY 3223 GGAAAGATACACTCACCGGAGAAAAGATTTCTGAAGAAAACCAACCATGCAATGA 3282
DB 3397 GGAAAGATACACTCACCGGAGAAAAGATTTCTGAAGAAAACCAACCATGCAATGA 3456
QY 3283 AGCAATGCTATTTCAATGGTCTCTTTTGTGAATGCAATTAATCAACAAAGGCTTTGATGA 3342
DB 3457 AGCAATGCTATTTCAATGGTCTCTTTTGTGAATGCAATTAATCAACAAAGGCTTTGATGA 3516
QY 3343 AAGGCATCGGTACATAGTGTATGTTTGGAGCTGGCATTTATTTGCTGAAAATCTTTC 3402
DB 3517 AAGGCATCGGTACATAGTGTATGTTTGGAGCTGGCATTTATTTGCTGAAAATCTTTC 3576
QY 3403 CAAAAGCAATCAATATGTATATGGAATTTGGAGAGGTACTGGGTGCTCCAGTTCAACAAGA 3462
DB 3577 CAAAAGCAATCAATATGTATATGGAATTTGGAGAGGTACTGGGTGCTCCAGTTCAACAAGA 3636
QY 3463 CAGATCTTTGTTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTC 3522
DB 3637 CAGATCTTTGTTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTC 3696
QY 3523 TTTTCTCGAGTTCAGTGCATCAATGCAATGCAATTCCTCTCCAGTCACTCAGTTCAC 3582
DB 3697 TTTTCTCGAGTTCAGTGCATCAATGCAATGCAATTCCTCTCCAGTCACTCAGTTCAC 3756
QY 3583 TGGTAGGCCAGTGTAAATGGCTAGCATTTAGCTGAATATGTTTACAGAGGAGAAACA 3642
DB 3757 TGGTAGGCCAGTGTAAATGGCTAGCATTTAGCTGAATATGTTTACAGAGGAGAAACA 3816
QY 3643 GCCTTATCTCGATATTTAAATTTACTTACCAGATTTATGAGGCTCAAGGTATGGTCGATGG 3702
DB 3817 GCCTTATCTCGATATTTAAATTTACTTACCAGATTTATGAGGCTCAAGGTATGGTCGATGG 3876
QY 3703 ATAATAGTTATTTTAAAGAACTAATTCACCTGAACCTTAAATCATCAACAGCAGCAGTGG 3762
DB 3877 ATAATAGTTATTTTAAAGAACTAATTCACCTGAACCTTAAATCATCAACAGCAGCAGTGG 3936

QY 3763 CCTCTAGCTTTTACTCCTTTGCTGTAAGAAAAA 3794
DB 3937 CCTCTAGCTTTTACTCCTTTGCTGTAAGAAAAA 3968
RESULT 10
AXI33694 LOCUS 3394 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0130987.
ACCESSION AXI33694
VERSION AXI33694.1 GI:14139712
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Luo,Y., Chan,E., Xu,X. and Huang,B.
TITLE Tankyrase h, compositions involved in the cell cycle and methods of use
JOURNAL Patent: WO 0130987-A 1 03-MAY-2001;
Rigel Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
source 1..3394
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1046 a 662 c 767 g 919 t
ORIGIN
Query Match 89.2%; Score 3386; DB 6; Length 3394;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3389; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 404 GGTTTTGGCGGAAAGACGCTAGTTGAATATTTCTTCAGAACTGTCGAAGTCCCAAGCA 463
DB 1 GGTTTTGGCGGAAAGACGCTAGTTGAATATTTCTTCAGAACTGTCGAAGTCCCAAGCA 60
QY 464 CGTGATGATGGGGCCCTATTCCCTCTTCATAATGCATGCTCTTTTGGTCTATGCTGAAGTA 523
DB 61 CGTGATGATGGGGCCCTATTCCCTTCATAATGCATGCTCTTTTGGTCTATGCTGAAGTA 120
QY 524 GTCAATCTCTTTTGGGACATGTGCAGACCCCAATGCTCGAGATTAATGGAATTTACT 583
DB 121 GTCAATCTCTTTTGGGACATGTGCAGACCCCAATGCTCGAGATTAATGGAATTTACT 180
QY 584 CCCTCCATCAAGCTGCAATTAAGGAAAGATTTGATGTTTGGCATTGCTGTTACAGCAT 643
DB 181 CCCTCCATCAAGCTGCAATTAAGGAAAGATTTGATGTTTGGCATTGCTGTTACAGCAT 240
QY 644 GGAGCTGAGCCCAACCTCCGAAATACAGATGGAAGGACAGCATTTGGATTAGCAGATCCA 703
DB 241 GGAGCTGAGCCCAACCTCCGAAATACAGATGGAAGGACAGCATTTGGATTAGCAGATCCA 300
QY 704 TCTGCCAAAGCAGTGTCTACTGTTGAATATAAGAAAGATGAACCTCTAGAAAGTCCAGG 763
DB 301 TCTGCCAAAGCAGTGTCTACTGTTGAATATAAGAAAGATGAACCTCTAGAAAGTCCAGG 360
QY 764 AGTGGCAATCAAGAAAAATGATGGCTCTACTCACACCATTAATGTCAACTGCCAGCA 823
DB 361 AGTGGCAATCAAGAAAAATGATGGCTCTACTCACACCATTAATGTCAACTGCCAGCA 420
QY 824 AGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 883
DB 421 AGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 480
QY 884 GTACAGTGTCTACTGCAACATGGAGCTGATGTCCTACTGCTAAAGATAAAGGTGATCTGGTA 943
DB 481 GTACAGTGTCTACTGCAACATGGAGCTGATGTCCTACTGCTAAAGATAAAGGTGATCTGGTA 540
QY 944 CCATTACAAATGCCTGTTCTTTATGGTCAATTTAGAGTAAGTGAACCTTTTGGTCAAGCAT 1003
DB 541 CCATTACAAATGCCTGTTCTTTATGGTCAATTTAGAGTAAGTGAACCTTTTGGTCAAGCAT 600

QY 1004 GTGGCTGTGTAATGCAATGGACTGTGGCAATTCACCTCCTTCATGAGCGACCTTCT 1063
Db 601 GGTGCTGTGTAATGCAATGGACTGTGGCAATTCACCTCCTTCATGAGCGACCTTCT 660
QY 1064 AAGAACAGGGTGAAGTATGTTCTCTTCCTTAAGTTATGGTGCAGACCAACACTGCTC 1123
Db 661 AAGAACAGGGTGAAGTATGTTCTCTCTTCCTTAAGTTATGGTGCAGACCAACACTGCTC 720
QY 1124 AATTGTCAACAATAAAGTCTATAGACTTGGCTCCACACACAGTAAAGAAAGATTA 1183
Db 721 AATTGTCAACAATAAAGTCTATAGACTTGGCTCCACACACAGTAAAGAAAGATTA 780
QY 1184 GCATATGAATTAAGGCCACTCGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTCGA 1243
Db 781 GCATATGAATTAAGGCCACTCGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTCGA 840
QY 1244 ATCAAAAAACATCTCTCTCGAATGTGTAATTTCAAGCATCCTCAACACATGAACA 1303
Db 841 ATCAAAAAACATCTCTCTCGAATGTGTAATTTCAAGCATCCTCAACACATGAACA 900
QY 1304 GCATTGCATTGTGCTGCTCATCTCCATATCCAAAAGAAAGCAAAATATGTGAACCTGTTG 1363
Db 901 GCATTGCATTGTGCTGCTCATCTCCATATCCAAAAGAAAGCAAAATATGTGAACCTGTTG 960
QY 1364 CTAAGAAAAGGAGCAACATCAATGAAAGACTAAGAAATCTTGACTCCTCTGCACGTG 1423
Db 961 CTAAGAAAAGGAGCAACATCAATGAAAGACTAAGAAATCTTGACTCCTCTGCACGTG 1020
QY 1424 GCATCTGAGAAAGCTCAATATCATGTGTTGAAGTAGTGGTGAACATCAAGCAAGGTT 1483
Db 1021 GCATCTGAGAAAGCTCAATATCATGTGTTGAAGTAGTGGTGAACATCAAGCAAGGTT 1080
QY 1484 AATGCTCTGGTAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATCTA 1543
Db 1081 AATGCTCTGGTAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATCTA 1140
QY 1544 CAAACCTCGCCGCTACTCCTGAGCTATGGGTGTGATCTCTAACATTAATATCCCTTCAGGC 1603
Db 1141 CAAACCTCGCCGCTACTCCTGAGCTATGGGTGTGATCTCTAACATTAATATCCCTTCAGGC 1200
QY 1604 TTACTGCTTTACAGATGGGAATGAAATGACAGCACTCCTCCAGAGGGTATCTCA 1663
Db 1201 TTACTGCTTTACAGATGGGAATGAAATGACAGCACTCCTCCAGAGGGTATCTCA 1260
QY 1664 TPAGTATATCAGAGGCAGACAGCAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGAA 1723
Db 1261 TPAGTATATCAGAGGCAGACAGCAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGAA 1320
QY 1724 ACTGTAAAAAACTGTGTACTGTTACAGAGTGTCAGCTGCAGAGACATTTGAAGGGGCTCAG 1783
Db 1321 ACTGTAAAAAACTGTGTACTGTTACAGAGTGTCAGCTGCAGAGACATTTGAAGGGGCTCAG 1380
QY 1784 TCTACACCACTTCATTTGACGCTGGGTATACAGAGTGTCCGTTGGTAATATCTGCTA 1843
Db 1381 TCTACACCACTTCATTTGACGCTGGGTATACAGAGTGTCCGTTGGTAATATCTGCTA 1440
QY 1844 CAGCATGAGCTGTGATGCTGCTAAAGATAAAGAGAGGCTTGTACCTTTGCACAAATGCA 1903
Db 1441 CAGCATGAGCTGTGATGCTGCTAAAGATAAAGAGAGGCTTGTACCTTTGCACAAATGCA 1500
QY 1904 TGTTCTTATGGACATATCAAGTTCCAGAACTCTTGTAAACATGGAGCAGTATGTTAT 1963
Db 1501 TGTTCTTATGGACATATCAAGTTCCAGAACTCTTGTAAACATGGAGCAGTATGTTAT 1560
QY 1964 GTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGACGAAAAAGGAAATATGAA 2023
Db 1561 GTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGACGAAAAAGGAAATATGAA 1620
QY 2024 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTCATCCAAAAAAGAGGATGGAAT 2083
Db 1621 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTCATCCAAAAAAGAGGATGGAAT 1680
QY 2084 ACTCCTTTGGATCTTTGTAAGATGGAGATACAGATATTCAGATCTGCTTAGGGAGNT 2143

Db 1681 ACTCCTTTGGATCTTTGTTAAAGATGAGATACAGATATTCATTTATCTGCTTAGGGAGAT 1740
QY 2144 GCAGCTTTGCTAGATGCTGCCAAGAGAGGTTCTTTAGCCAGAGTGAAGAGTTGCTTCT 2203
Db 1741 GCAGCTTTGCTAGATGCTGCCAAGAGAGGTTCTTTAGCCAGAGTGAAGAGTTGCTTCT 1800
QY 2204 CCTGATAATGTAATTTGCCGCGATACCCCAAGCAGACATTCACACCTTTACATTTAGCA 2263
Db 1801 CCTGATAATGTAATTTGCCGCGATACCCCAAGCAGACATTCACACCTTTACATTTAGCA 1860
QY 2264 GCTGTTATATAATTTAGAACTTGCAGAGTATTTGTTTACAACAGGAGCTGATGTGAAT 2323
Db 1861 GCTGTTATATAATTTAGAACTTGCAGAGTATTTGTTTACAACAGGAGCTGATGTGAAT 1920
QY 2324 GCCCAGACAAAGGAGGACTTATCTCTTACATAATGCAGCATCTTACGGGATGTAGAT 2383
Db 1921 GCCCAGACAAAGGAGGACTTATCTCTTACATAATGCAGCATCTTACGGGATGTAGAT 1980
QY 2384 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGACAAATGGGCTTTC 2443
Db 1981 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGACAAATGGGCTTTC 2040
QY 2444 ACACCTTTGCACGAGCAGCCCAAAAGGAGCAACACAGCTTTTGTGCTTAGCC 2503
Db 2041 ACACCTTTGCACGAGCAGCCCAAAAGGAGCAACACAGCTTTTGTGCTTAGCC 2100
QY 2504 CATGGAGCTGACCCGACTCTTAAATAATCAGGAAGACAAACACCTTTAGATTTAGTTCA 2563
Db 2101 CATGGAGCTGACCCGACTCTTAAATAATCAGGAAGACAAACACCTTTAGATTTAGTTCA 2160
QY 2564 GCGGATGATGTACAGGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTGCCCTCTTGT 2623
Db 2161 GCGGATGATGTACAGGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTGCCCTCTTGT 2220
QY 2624 TACAGCCTCAAGTCTCAATGGTGTGAGAACCCAGAGCCACTGCAGATGCTCTCT 2683
Db 2221 TACAGCCTCAAGTCTCAATGGTGTGAGAACCCAGAGCCACTGCAGATGCTCTCT 2280
QY 2684 TCAGTCTCATCTAGCCCATCAAGCCTTCTGACAGCAGCAGCTTTCAGAACCTTATCTGGG 2743
Db 2281 TCAGTCTCATCTAGCCCATCAAGCCTTCTGACAGCAGCAGCTTTCAGAACCTTATCTGGG 2340
QY 2744 AGTTTTTCAAGACTGCTCTCAGTATAGTTCAAGTGGAAACAGAGGCTTCCAGTTTG 2803
Db 2341 AGTTTTTCAAGACTGCTCTCAGTATAGTTCAAGTGGAAACAGAGGCTTCCAGTTTG 2400
QY 2804 GAGAAAAGGAGTTCCAGGAGTAGATTTTACATTAACCTCAATTCGTAAAGAACTCTTGA 2863
Db 2401 GAGAAAAGGAGTTCCAGGAGTAGATTTTACATTAACCTCAATTCGTAAAGAACTCTTGA 2460
QY 2864 CTTGAGCACCTAATCGATATATTTGAGAGAACAGATCACCTTTGGATGTATTAGTTGAG 2923
Db 2461 CTTGAGCACCTAATCGATATATTTGAGAGAACAGATCACCTTTGGATGTATTAGTTGAG 2520
QY 2924 ATGGGCAACAGGAGCTGAAGGAGATTTGAATCAATGTTTATGGACATAGGCAAACTA 2983
Db 2521 ATGGGCAACAGGAGCTGAAGGAGATTTGAATCAATGTTTATGGACATAGGCAAACTA 2580
QY 2984 ATTAAGGAGTGCAGAGACTTATCTCCGACAAAGGCTTTAACCCATATTTTAACTTTG 3043
Db 2581 ATTAAGGAGTGCAGAGACTTATCTCCGACAAAGGCTTTAACCCATATTTTAACTTTG 2640
QY 3044 AACACCTCTGTTAGTGGAACTTCTTATAGATCTCTCTCTCTGATGATAAGAGTTTCAG 3103
Db 2641 AACACCTCTGTTAGTGGAACTTCTTATAGATCTCTCTCTCTGATGATAAGAGTTTCAG 2700
QY 3104 TCTGTGGAGGAAGATGCAAAAGTACAGTTCAGAGCAGACAGAGTGGAGTCAATGCAGT 3163
Db 2701 TCTGTGGAGGAAGATGCAAAAGTACAGTTCAGAGCAGACAGAGTGGAGTCAATGCAGT 2760
QY 3164 GGAATCTTCAACAGATACAAATTTCTCAAGATTCAGAGGTTTGTAAACAGAACTATCG 3223

Db 2761 GGAATCTTCAACAGATACAAATATTTCTCAAGATTTCAGAAAGTTTGTAAACAAGAACTATGG 2820
Qy 3224 GAAAGATACACTCACCAGAGAGAAAGAGTTTCTTGAAGAAACACACACCATGCCAATGAA 3283
Db 2821 GAAAGATACACTCACCAGAGAGAAAGAGTTTCTTGAAGAAACACACACCATGCCAATGAA 2880
Qy 3284 GGAATGCTATTTTCATGGGTCTCTTTTGTGAATGCAATTTATCCAAAGGCTTTTGATGAA 3343
Db 2881 GGAATGCTATTTTCATGGGTCTCTTTTGTGAATGCAATTTATCCAAAGGCTTTTGATGAA 2940
Qy 3344 AGGCATGGGTACATAGTGGTATGTTTGGAGCTGGCAATTTATTTTGTGAAACCTCTTCC 3403
Db 2941 AGGCATGGGTACATAGTGGTATGTTTGGAGCTGGCAATTTATTTTGTGAAACCTCTTCC 3000
Qy 3404 AAAAGCAATCAATATGATATGGAATTTGGAGGAGGTACTGGGTGCCAGTTTCACAAAGAC 3463
Db 3001 AAAAGCAATCAATATGATATGGAATTTGGAGGAGGTACTGGGTGCCAGTTTCACAAAGAC 3060
Qy 3464 AGATCTTGTGTACATTTGGCACAGGAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTCT 3523
Db 3061 AGATCTTGTGTACATTTGGCACAGGAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTCT 3120
Qy 3524 TTCTGCGAGTTTCAGTGAATGAAATGGCAATCTCTCCAGGTCACTCAGTCACT 3583
Db 3121 TTCTGCGAGTTTCAGTGAATGAAATGGCAATCTCTCCAGGTCACTCAGTCACT 3180
Qy 3584 GGTAGGCCAGGTAAATGGCTAGCATTAGCTGAATATGTTTATTCAGAGGAGACAG 3643
Db 3181 GGTAGGCCAGGTAAATGGCTAGCATTAGCTGAATATGTTTATTCAGAGGAGACAG 3240
Qy 3644 GCTTATCTCTGAGTATTTAATTAATTTACCAGATTTATGAGGCTTGAAGGTATGTCGATGGA 3703
Db 3241 GCTTATCTCTGAGTATTTAATTAATTTACCAGATTTATGAGGCTTGAAGGTATGTCGATGGA 3300
Qy 3704 TAAATAGTATTTTAAAGAACTAATTCACAGAACTTAAATCATCAAGCAGAGTGGC 3763
Db 3301 TAAATAGTATTTTAAAGAACTAATTCACAGAACTTAAATCATCAAGCAGAGTGGC 3360
Qy 3764 CTTCTAGGTTTACTCTTCTTGGTGAAGAAAAA 3797
Db 3361 CTTCTAGGTTTACTCTTCTTGGTGAAGAAAAA 3394

RESULT 11
AX076213
LOCUS AX076213 4512 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 3 from Patent WO0104326.
ACCESSION AX076213
VERSION AX076213.1 GI:12710838
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4512)
AUTHORS Berthelsen, J., Toma, S. and Isacchi, A.
TITLE Tankyrase homolog protein (thp), nucleic acids, and methods related to the same
JOURNAL Patent: WO 0104326-A 3 18-JAN-2001;
PHARMACIA & UPJOHN S.p.A. (IT)
FEATURES Location/Qualifiers
source 1..4512
/organism="Homo sapiens"
/db_xref="taxon:9606"
1124
unsure /note="n is any base"
2672
unsure /note="n is any base"
BASE COUNT 1356 a 878 c 1002 g 1268 t 8 others
ORIGIN
Query Match 89.1%; Score 3382.8; DB 6; Length 4512;
Best Local Similarity 99.7%; Pred. No. 0;

Matches 3381; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
Qy 403 AGGTTTGGCGGAAAGACGCTAGTTGAATATTTGCTTCAAGATGGTCAAGTGTCCAAGC 462
Db 220 AGGTTTGGCGGAAAGACGCTAGTTGAATATTTGCTTCAAGATGGTCAAGTGTCCAAGC 279
Qy 463 ACCTGATGATGGGGGCTTATTCCTCTTTCATAATGATGCTCTTTTGGTCACTGCTGAAGT 522
Db 280 ACCTGATGATGGGGGCTTATTCCTCTTTCATAATGATGCTCTTTTGGTCACTGCTGAAGT 339
Qy 523 AGTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTTATAC 582
Db 340 AGTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTTATAC 399
Qy 583 TCCTCTCCATGAAGCTGCAATTTAAAGGAAAGATTGATTTGTCATTTGCTGTTTACAGCA 642
Db 400 TCCTCTCCATGAAGCTGCAATTTAAAGGAAAGATTGATTTGTCATTTGCTGTTTACAGCA 459
Qy 643 TGGAGCTGAGCCCAACCATCGAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC 702
Db 460 TGGAGCTGAGCCCAACCATCGAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC 519
Qy 703 ATCTGCCAAAGCAGTCTTACTGCTGAATATAGAAAGATGAACCTCTTAGAAAAGTCCAG 762
Db 520 ATCTGCCAAAGCAGTCTTACTGCTGAATATAGAAAGATGAACCTCTTAGAAAAGTCCAG 579
Qy 763 GAGTGGCAATGAAGAAAAATGATGGCTCTACTCACACCATTTAAATGTCAACTGCCACGC 822
Db 580 GAGTGGCAATGAAGAAAAATGATGGCTCTACTCACACCATTTAAATGTCAACTGCCACGC 639
Qy 823 AAGTGTGGCAGAAAGTCAACTCCATTTACATTTGGCAGCAGGATATTAACAGAGTAAAGAT 882
Db 640 AAGTGTGGCAGAAAGTCAACTCCATTTACATTTGGCAGCAGGATATTAACAGAGTAAAGAT 699
Qy 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGCTCATGCTTAAAGATAAAGTGATCTGGT 942
Db 700 TGTACAGCTGTTACTGCAACATGGAGCTGATGCTCATGCTTAAAGATAAAGTGATCTGGT 759
Qy 943 ACCATTACACAATGCCTGTTCTTTATGCTCATTTATGAAGTAACCTTTTGGTCAAGCA 1002
Db 760 ACCATTACACAATGCCTGTTCTTTATGCTCATTTATGAAGTAACCTTTTGGTCAAGCA 819
Qy 1003 TGGTGCCTGTGTAATGCAATGGACCTTGGCAATTCACCTCTTCATGAGGAGCTTC 1062
Db 820 TGGTGCCTGTGTAATGCAATGGACCTTGGCAATTCACCTCTTCATGAGGAGCTTC 879
Qy 1063 TAAGAACAGGCTTGAAGTATGTTCTCTCTTAAAGTTATGGTGCAGACCCCAACACTGCT 1122
Db 880 TAAGAACAGGCTTGAAGTATGTTCTCTCTTAAAGTTATGGTGCAGACCCCAACACTGCT 939
Qy 1123 CAATTGTCAACAATAAAGTGTATAGACTTGGCTCCCAACACACAGTTAAAGAAAGATT 1182
Db 940 CAATTGTCAACAATAAAGTGTATAGACTTGGCTCCCAACACACAGTTAAAGAAAGATT 999
Qy 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGGAGAGCTGATTTACTCG 1242
Db 1000 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGGAGAGCTGATTTACTCG 1059
Qy 1243 AATCAAAAAACATCTCTCTCTGGAATGTTTCAAGCATCTCAAAACACATGAAC 1302
Db 1060 AATCAAAAAACATCTCTCTCTGGAATGTTTCAAGCATCTCAAAACACATGAAC 1119
Qy 1303 AGCATTTGATTTGCTGCTGCATCTCCATATCCCAAGAAAGCAAAATATGTGAACCTGT 1362
Db 1120 AGCATTTGATTTGCTGCTGCATCTCCATATCCCAAGAAAGCAAAATATGTGAACCTGT 1179
Qy 1363 GCTAAGAAAGAGGACAAACATCAATGAAGAACTTAAAGAAATCTTCTGCTGACCT 1422
Db 1180 GCTAAGAAAGAGGACAAACATCAATGAAGAACTTAAAGAAATCTTCTGCTGACCT 1239
Qy 1423 GGCATCTCAGAAAGCTCATATGATGTTTGAAGTAGTGTGAACATGAAGCAAGGT 1482
Db 1240 GGCATCTCAGAAAGCTCATATGATGTTTGAAGTAGTGTGAACATGAAGCAAGGT 1299

QY 1483 TAATGCTCTGATGAATCTTGGTCAGACTTCTCTACAGAGCTGCAATATTTGTTGCTCATCT 1542
DB 1300 TAATGCTCTGATGAATCTTGGTCAGACTTCTCTACAGAGCTGCAATATTTGTTGCTCATCT 1359
QY 1543 ACAAACTGCGCCCTACTCTGAGCTATGCGTGTGATGCTTAAATATATATCCCTTCAGGG 1602
DB 1360 ACAAACTGCGCCCTACTCTGAGCTATGCGTGTGATGCTTAAATATATATCCCTTCAGGG 1419
QY 1603 CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCCAAGAGGATCTTC 1662
DB 1420 CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCCAAGAGGATCTTC 1479
QY 1663 ATTAGGTAATTCAGAGCGACAGACAAATTTGCTGGAAGCTGCAAGAGCTGAGATGTGGA 1722
DB 1480 ATTAGGTAATTCAGAGCGACAGACAAATTTGCTGGAAGCTGCAAGAGCTGAGATGTGGA 1539
QY 1723 AACTGTAATAAANAAGTGTACTGTTACAGAGTGTCAACTGTCAGAGACATTTGAAGGGGTCA 1782
DB 1540 AACTGTAATAAANAAGTGTACTGTTACAGAGTGTCAACTGTCAGAGACATTTGAAGGGGTCA 1599
QY 1783 GTCTACACCACTTCATTTTGACGTGGGTATAACAGAGTGTCCGTGTTGGAATATCTGCT 1842
DB 1600 GTCTACACCACTTCATTTTGACGTGGGTATAACAGAGTGTCCGTGTTGGAATATCTGCT 1659
QY 1843 ACAGCATGGAGCTGATGTGATGCTTAAAGATAAAGGAGGCTTTGTACCTTTGCAACAATGC 1902
DB 1660 ACAGCATGGAGCTGATGTGATGCTTAAAGATAAAGGAGGCTTTGTACCTTTGCAACAATGC 1719
QY 1903 ATGTTCTTATGGACATTTCAAGTTCGAGAACTCTTGTAAACATGGAGCAGTAGTTAA 1962
DB 1720 ATGTTCTTATGGACATTTCAAGTTCGAGAACTCTTGTAAACATGGAGCAGTAGTTAA 1779
QY 1963 TGTAGCTGATTTTGGAAATTTTACACTTTACATGAAGCAGCAGCAAAAGGAAATATGA 2022
DB 1780 TGTAGCTGATTTTGGAAATTTTACACTTTACATGAAGCAGCAGCAAAAGGAAATATGA 1839
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTCAGACCTTACCAAAAAAAGGAGTGGAAA 2082
DB 1840 AATTTGCAAACTTCTGCTCCAGCATGGTCAGACCTTACCAAAAAAAGGAGTGGAAA 1899
QY 2083 TACTCCTTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTCTTAGGGGAGA 2142
DB 1900 TACTCCTTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTCTTAGGGGAGA 1959
QY 2143 TGCAGCTTTGCTAGTGTGCTCCAGAGGTTGTTTTCAGCAGAGTGAAGAGTGTCTTC 2202
DB 1960 TGCAGCTTTGCTAGTGTGCTCCAGAGGTTGTTTTCAGCAGAGTGAAGAGTGTCTTC 2019
QY 2203 TCTGATAATGTAATTTGCGCGGATACCAAGCAGACATTTCAACACCTTTACATTTAGC 2262
DB 2020 TCTGATAATGTAATTTGCGCGGATACCAAGCAGACATTTCAACACCTTTACATTTAGC 2079
QY 2263 AGCTGTTTATAATTTAGAAGTTCAGAGATTTGTTTACACAGGAGCTGATGTGAA 2322
DB 2080 AGCTGTTTATAATTTAGAAGTTCAGAGATTTGTTTACACAGGAGCTGATGTGAA 2139
QY 2323 TGCCCAAGACAAGAGGAGCTTATTCCTTTACATATATGACAGATCTTACGGGCATGTAGA 2382
DB 2140 TGCCCAAGACAAGAGGAGCTTATTCCTTTACATATATGACAGATCTTACGGGCATGTAGA 2199
QY 2383 TGTAGCAGCTCTACTAATAAAGTATAATGATGCTGTCATGTCAGGAGCAAAATGGGCTTT 2442
DB 2200 TGTAGCAGCTCTACTAATAAAGTATAATGATGCTGTCATGTCAGGAGCAAAATGGGCTTT 2259
QY 2443 CACACCTTTGACAGAGCAGCCCAAAAGGAGCAACAGAGCTTTGCTTTGCTTCTAGC 2502
DB 2260 CACACCTTTGACAGAGCAGCCCAAAAGGAGCAACAGAGCTTTGCTTTGCTTCTAGC 2319
QY 2503 CCATGAGCTGACCGACTCTTTAAATAACAGGAAGGACAAACCTTTAGATTTAGTTTC 2562
DB 2320 CCATGAGCTGACCGACTCTTTAAATAACAGGAAGGACAAACCTTTAGATTTAGTTTC 2379

QY 2563 AGCGATGATGTACAGCGCTCTTCTGACAGAGCATGCCCATCTGCTCTGCCCTCTTG 2622
DB 2380 AGCGATGATGTACAGCGCTCTTCTGACAGAGCATGCCCATCTGCTCTGCCCTCTTG 2439
QY 2623 TTACAAGCCTCAAGTGTCTCAATGGTGTGAGAGCCCAAGAGCCACTGAGATGCTCTCTC 2682
DB 2440 TTACAAGCCTCAAGTGTCTCAATGGTGTGAGAGCCCAAGAGCCACTGAGATGCTCTCTC 2499
QY 2683 TTCAGGTCCATCTAGCCCATCAAGCTTTCTGACAGCCAGAGCTTTGACAACTTATCTGG 2742
DB 2500 TTCAGGTCCATCTAGCCCATCAAGCTTTCTGACAGCCAGAGCTTTGACAACTTATCTGG 2559
QY 2743 GAGTTTTTTTCAGAACTGTCTTCAAGTGTAGTTTCAAGTGTGAGAGGGTCTCCAGTTT 2802
DB 2560 GAGTTTTTTTCAGAACTGTCTTCAAGTGTAGTTTCAAGTGTGAGAGGGTCTCTCCAGTTT 2619
QY 2803 GGAGAAAAGGAGGTTCCAGGAGTAGATTTTACATAACTCAATTCSTPAAGAACTTTTGG 2862
DB 2620 GGAGAAAAGGAGGTTCCAGGAGTAGATTTTACATAACTCAATTCSTPAAGAACTTTTGG 2679
QY 2863 ACTTACAGCAGCTTAATGGATATATTTGAGAGAGACAGATCACTTTGGATGTATTAGTTGA 2922
DB 2680 ACTTACAGCAGCTTAATGGATATATTTGAGAGAGACAGATCACTTTGGATGTATTAGTTGA 2739
QY 2923 GATGGGACACAGGAGCTCAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACT 2982
DB 2740 GATGGGACACAGGAGCTCAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACT 2799
QY 2983 AATTAAAGAGTTCGAGAGCTTTATCTCCGGACAAAGGTTCTTAACCCATATTTAACTTT 3042
DB 2800 AATTAAAGAGTTCGAGAGCTTTATCTCCGGACAAAGGTTCTTAACCCATATTTAACTTT 2859
QY 3043 GAACACCTCTGTTAGTGAACAATTTTATAGATCTGCTCTCTGATGATAAAGAGTTTCA 3102
DB 2860 GAACACCTCTGTTAGTGAACAATTTTATAGATCTGCTCTCTGATGATAAAGAGTTTCA 2919
QY 3103 GTCTGTGGAGGAGAGATGCAAAAGTTCAGAGTTCGAGAGCAGAGATGGAGTCTATGCAGG 3162
DB 2920 GTCTGTGGAGGAGAGATGCAAAAGTTCAGAGTTCGAGAGCAGAGATGGAGTCTATGCAGG 2979
QY 3163 TGGATCTTTCAACAGATACATATTTCTCAAGATTCTCAAGATTCTCAAGATTCTCAAGATTCT 3222
DB 2980 TGGATCTTTCAACAGATACATATTTCTCAAGATTCTCAAGATTCTCAAGATTCTCAAGATTCT 3039
QY 3223 GGAAGATACACTCACCGGAGAAAGAGTTCTTGAAGAAACCAACCATGCAATGA 3282
DB 3040 GGAAGATACACTCACCGGAGAAAGAGTTCTTGAAGAAACCAACCATGCAATGA 3099
QY 3283 ACGAATGCTATTTTCATGGTCTCCTTTTGTGAATGCAATTTATCCAAAGAGCTTTTGATGA 3342
DB 3100 ACGAATGCTATTTTCATGGTCTCCTTTTGTGAATGCAATTTATCCAAAGAGCTTTTGATGA 3159
QY 3343 AAGCATGCTGATACATAGTGTATGTTTGGAGCTGGCATTTATTTTCTGCTGAAACCTTTC 3402
DB 3160 AAGCATGCTGATACATAGTGTATGTTTGGAGCTGGCATTTATTTTCTGCTGAAACCTTTC 3219
QY 3403 CAAAAGCAATCAATATGATGTAATGGAATTTGAGGAGTACTGGGTGTCAGCTTCCAAAGA 3462
DB 3220 CAAAAGCAATCAATATGATGTAATGGAATTTGAGGAGTACTGGGTGTCAGCTTCCAAAGA 3279
QY 3463 CAGATCTTTGATACATTTGCCACAGGAGCTGCTCTTTTGGCCGGTAACTTTGGGAAGTC 3522
DB 3280 CAGATCTTTGATACATTTGCCACAGGAGCTGCTCTTTTGGCCGGTAACTTTGGGAAGTC 3339
QY 3523 TTTTCTCCAGTTCAGTCAATGAAATGGCAGATTTCTTCTCCAGGTCATCTCAGTTCAC 3582
DB 3340 TTTTCTCCAGTTCAGTCAATGAAATGGCAGATTTCTTCTCCAGGTCATCTCAGTTCAC 3399
QY 3583 TGTAGGCGCCAGTGTAAATGGCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAA 3642
DB 3400 TGTAGGCGCCAGTGTAAATGGCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAA 3459
QY 3643 GGCTTATCTTGAGTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3702


```

Db 3736 AGCTTATCTGAGTATTTAATTACTTACACAGATTATGAGCGCTGAGGCTATGTCGATG 3795
QY 3702 GATAAATAGTATTTTAAAGAACTAATTCACATGAACCTTAAATCATCAAGCAGCAGTG 3761
Db 3796 GATAAATAGTATTTTAAAGAACTAATTCACATGAACCTTAAATCATCAAGCAGCAGTG 3855
QY 3762 GCCTCTACGTTTACTCCTTGTGCTGAAAAAAA 3794
Db 3856 GCCTCTACGTTTACTCCTTGTGCTGAAAAAAA 3888

RESULT 13
AF342982
LOCUS Homo sapiens tankyrase 2 mRNA, complete cds.
DEFINITION AF342982
ACCESSION AF342982.1 GI:13430364
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3501)
Kaminker,P.G., Kim,S.H., Taylor,R.D., Zebajadian,Y., Funk,W.D.,
Morin,G.B., Yaswen,P. and Campisi,J.
TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes
rapid induction of cell death upon overexpression
J. Biol. Chem. 276 (38), 35891-35899 (2001)
21443728
PUBMED 11454873
2 (bases 1 to 3501)
Kaminker,P.G., Campisi,J., Kim,S.H., Yaswen,P. and Morin,G.
Direct Submission
Submitted (25-JAN-2001) Life Sciences, Lawrence Berkeley National
Labs, 1 Cyclotron Rd 84-171, Berkeley, CA 94720, USA
FEATURES
source
1..3501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q23-q24"
1..3501
/note="TANK2; telomere-associated poly(ADP-ribose)
polymerase; contains ankyrin domain, sterile alpha motif
and PARP catalytic domain; similar to TNKS"
/codon_start=1
/product="tankyrase 2"
/protein_id="AAK25811.1"
/db_xref="GI:13430365"
/translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
TPEKNSRDTAGRKSTPLHFAAGFKRDVVEYLQNGANVOARDGGLIPLHNACSPG
HAEVNLRLRHGADPNARDNWNYPLEAAIKGKIDVCIVLQHCAPTIRNTDGPRTA
LDADPSAKAVLTGEYKKDELLESARSNEKMMALLTPLNVNCHASDGRKSTPLHLA
AGYNRVKIOLLLOHGDVHAKDKDGLVPLHNACSYGHEVEITLLVHRGACVNMDLW
LQAREADVTIRKHLKLSLEMYNEKHPOTHEALHCAASPYPKRQICELLIRKGAN
INEXTKEITPLHVAASEKNDVVEVVKHEAKVNDNLGQTSLSHRAAYCGHLQTCR
LLSYGCDPNTIISLQGTALQMGNEVQQLQEGISLGNSEADRLLEAAKAGDVETV
KKLCTVQSVNCRDETRQSTPLHFAAGYNRVVEVYLLQHCADVHAKDKGLVPLHNA
CSYGHVEAEVLLVKGAVNVADLWKPTPLHEAAKKGKYEICKLLLOHGDPTKKNKD
GNTPLDLVKDGDIDQLRGAALLDAARKKGLARVKLSSPDNVNCRDQGRHSTP
LHLAAGYNLEVBAYLIQHDADVNAQDKGLIPLHNAASYGHVDVAALLKYACVNA
TDKWAFTPLHFAAQKQRTQCALLAHGADPTLKNQBOGTPLDLSADDDYSALLTAA
PPSALPCYKQVQLNGVRSPGATLAHSGSPSSLSAASSLDNLSCSPSELSSVVS
SSGTGAGSSLEKKEVPGVDESITQFVNLGLHLMIDIFEREQITLDLVEMGHKELKE
IGINAYGHRHKLKGVRLISGQGLNPYTLNTSGSGTILIDLSPODKFEQSVSEEM
QSTVEHRDGHGAGIENRYNLIKQVCKNKLMEYTHRRKEVSEENHNHNMELF
HGSBFWNAIHKGEHRHAYIGGMFGAGIYFAENSSKNQYVYVIGSGTGCPCVHKDRS
CYCHROLLFCVRLTGRKSLQFSAMKMAHSPPHSHSVTGRPSVNLALAEVVIYRGEQ
AYPELYTIQIMRPEGMVDG"
1044 a 707 c 840 g 909 t 1 others
BASE COUNT
ORIGIN

```

```

Query Match      86.9%; Score 3300.4; DB 9; Length 3501;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3301; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 403 AGGTTTGGCGGAAAGAGCTAGTTCAATATTTGCTTCAGAATGGTGAAGTCCCAAGC 462
Db 198 AGGTTTGGCGGAAAGAGCTAGTTGAATATTTGCTTCAGAATGGTGAAGTCCCAAGC 257
QY 463 ACCTGATGATGGGGCGCTTATTCCTCTCATAAATCATGCTCTTTGGTCATGCTGAAGT 522
Db 258 ACCTGATGATGGGGCGCTTATTCCTCTCATAAATCATGCTCTTTGGTCATGCTGAAGT 317
QY 523 AGTCAATCTCCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATATATAC 582
Db 318 AGTCAATCTCCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATATATAC 377
QY 583 TCCTCTCCATGAAGCTGCAATTTAAAGGAAAGATTGATGCTTTTGCATTTGCTGTACAGCA 642
Db 378 TCCTCTCCATGAAGCTGCAATTTAAAGGAAAGATTGATGCTTTGCTGTGTACAGCA 437
QY 643 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGSACAGCATTTGGATTTAGCAGATCC 702
Db 438 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGSACAGCATTTGGATTTAGCAGATCC 497
QY 703 ATCTGCCAAAGCAGTGGCTTACTGCTGAATATAAGAAAGATGAATCTTTAGAAAAGTGCAG 762
Db 498 ATCTGCCAAAGCAGTGGCTTACTGCTGAATATAAGAAAGATGAATCTTTAGAAAAGTGCAG 557
QY 763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAATGTCACATGCCACGC 822
Db 558 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAATGTCACATGCCACGC 617
QY 823 AAGTATGCCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATACAGAGTAAAGAT 882
Db 618 AAGTATGCCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATACAGAGTAAAGAT 677
QY 883 TGTACAGCTGTACTGCAACATGGAGCTGATGTCATGCTAAAGATAAAGTGATCTGGT 942
Db 678 TGTACAGCTGTACTGCAACATGGAGCTGATGTCATGCTAAAGATAAAGTGATCTGGT 737
QY 943 ACCATTACACAATGCCTGCTTCTTATGGTCAATTATGAAGTAATCACTTTTGGTCAAGCA 1002
Db 738 ACCATTACACAATGCCTGCTTCTTATGGTCAATTATGAAGTAATCACTTTTGGTCAAGCA 797
QY 1003 TGGTGCCTGTCTAAATGCAATGGACTTGTGCAATTCACCTCTTCATGAGGAGCTTC 1062
Db 798 TGGTGCCTGTCTAAATGCAATGGACTTGTGCAATTCACCTCTTCATGAGGAGCTTC 857
QY 1063 TAAGAACAGGCTTGAAGTATGCTCTCTCTTAAGTTATGGTCGACACCCACACATGCT 1122
Db 858 TAAGAACAGGCTTGAAGTATGCTCTCTCTTAAGTTATGGTCGACACCCACACATGCT 917
QY 1123 CAATTGTCAATAAAGTGCATATAGACTTTGGCTCCACACACAGTTAAAGAAAGATT 1182
Db 918 CAATTGTCAATAAAGTGCATATAGACTTTGGCTCCACACACAGTTAAAGAAAGATT 977
QY 1183 AGCATATGAATTTAAAGGCCACTCGTTGTCGCAAGCTGCGAGAGAACTGATGTTACTCG 1242
Db 978 AGCATATGAATTTAAAGGCCACTCGTTGTCGCAAGCTGCGAGAGAACTGATGTTACTCG 1037
QY 1243 AATCAAAAACATCTCTCTGGAATTTTCAAGCTCCTCAACACATGAAC 1302
Db 1038 AATCAAAAACATCTCTCTGGAATTTTCAAGCTCCTCAACACATGAAC 1097
QY 1303 AGCATTCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAAATATGTCAACTGTT 1362
Db 1098 AGCATTCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAAATATGTCAACTGTT 1157
QY 1363 GCTAAGAAAAGGACCAACATCAATGAAAAGCTAAAGAAATTTTGTACTCCTCTGCACGT 1422
Db 1158 GCTAAGAAAAGGACCAACATCAATGAAAAGCTAAAGAAATTTTGTACTCCTCTGCACGT 1217
QY 1423 GGCATCTGAGAAAGCTCATATATGATGTTGTTGAAGTAGTGTGTAACATGAACAAAGGT 1482

```

|||||
Db 1218 GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGTAACATGAAGCAAGGT 1277
|||||
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTCTCTACAGAGAGCTGCATATGTTGGTCACT 1542
|||||
Db 1278 TAATGCTCTGGATAATCTTGGTCAGACTCTCTACAGAGAGCTGCATATGTTGGTCACT 1337
|||||
Qy 1543 ACAAACTGCGCCCTACTCTGAGCTATGGGTGTCATCCTAACATATATCCCTTCAGGG 1602
|||||
Db 1338 ACAAACTGCGCCCTACTCTGAGCTATGGGTGTCATCCTAACATATATCCCTTCAGGG 1397
|||||
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAATGATCAGCAACTCTCTCAAGAGGGTATCTC 1662
|||||
Db 1398 CTTTACTGCTTTACAGATGGGAAATGAAATGATCAGCAACTCTCTCAAGAGGGTATCTC 1457
|||||
Qy 1663 ATTAGGTAAATCAGAGGACAGACAAATGCTGGAAGCTGCAAGAGCTGGAGATGTCGA 1722
|||||
Db 1458 ATTAGGTAAATCAGAGGACAGACAAATGCTGGAAGCTGCAAGAGCTGGAGATGTCGA 1517
|||||
Qy 1723 AACTGTAAAAAATCTGTACTGTTTCAGAGTCTCAACTGCAGAGACATTTGAAGGGCTCA 1782
|||||
Db 1518 AACTGTAAAAAATCTGTACTGTTTCAGAGTCTCAACTGCAGAGACATTTGAAGGGCTCA 1577
|||||
Qy 1783 GTCTACACACTTCATTTTGCAGCTGGGTATTAACAGAGTGTCCGTGGTGGAAATATCTGCT 1842
|||||
Db 1578 GTCTACACACTTCATTTTGCAGCTGGGTATTAACAGAGTGTCCGTGGTGGAAATATCTGCT 1637
|||||
Qy 1843 ACAGCATGGAGCTGATGTCATGCTTAAGATTAAGAGAGGCTGTACCTTTGCACAATGC 1902
|||||
Db 1638 ACAGCATGGAGCTGATGTCATGCTTAAGATTAAGAGAGGCTGTACCTTTGCACAATGC 1697
|||||
Qy 1903 ATGCTCTTATGGACATTTATGAAGTTTCGAACTTCTGTAAACATGGAGCAGTAGTTAA 1962
|||||
Db 1698 ATGCTCTTATGGACATTTATGAAGTTTCGAACTTCTGTAAACATGGAGCAGTAGTTAA 1757
|||||
Qy 1963 TGPAGCTGATTTATGGAATTTTACACCTTTACATGAAGCAGCAGCAAAAAGGAAATATGA 2022
|||||
Db 1758 TGPAGCTGATTTATGGAATTTTACACCTTTACATGAAGCAGCAGCAAAAAGGAAATATGA 1817
|||||
Qy 2023 AATTTGCAAACTCTGCTCCAGCATGGTCAGACCTTACCAAGGAGGATGGAAA 2082
|||||
Db 1818 AATTTGCAAACTCTGCTCCAGCATGGTCAGACCTTACCAAGGAGGATGGAAA 1877
|||||
Qy 2083 TACTCCTTTTGGATCTGTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 2142
|||||
Db 1878 TACTCCTTTTGGATCTGTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 1937
|||||
Qy 2143 TGCAGCTTTGCTAGCTGCTGCCAAGAGGGTTGTTTACAGAGTGAAGAGTTGCTTTC 2202
|||||
Db 1938 TGCAGCTTTGCTAGCTGCTGCCAAGAGGGTTGTTTACAGAGTGAAGAGTTGCTTTC 1997
|||||
Qy 2203 TCCGTGAATTAATGTAATTTGCCGCGATACCCAGGCAGACATTTCAACACCTTTACATTTAGC 2262
|||||
Db 1998 TCCGTGAATTAATGTAATTTGCCGCGATACCCAGGCAGACATTTCAACACCTTTACATTTAGC 2057
|||||
Qy 2263 AGCTGTTTATAAATTTAGAAAGTTGCAGAGTATTTGTTTACAAACAGGAGCTGATGTGAA 2322
|||||
Db 2058 AGCTGTTTATAAATTTAGAAAGTTGCAGAGTATTTGTTTACAAACAGGAGCTGATGTGAA 2117
|||||
Qy 2323 TGGCCAGACAAAGGAGGACTTATCTCTTTACATTAATGTCAGCATCTTACGGCATGTAGA 2382
|||||
Db 2118 TGGCCAGACAAAGGAGGACTTATCTCTTTACATTAATGTCAGCATCTTACGGCATGTAGA 2177
|||||
Qy 2383 TGTAGCAGCTCTACTAATAAGTATATAATGTCATGTGTCATGTCCTCAAGGACAAATGGGCTTT 2442
|||||
Db 2178 TGTAGCAGCTCTACTAATAAGTATATAATGTCATGTGTCATGTCCTCAAGGACAAATGGGCTTT 2237
|||||
Qy 2443 CACACCTTTTGCAGAGAGAGCCCAAAAAGGAGCAGCAACAGCTTTGCTTTGTTGCTAGC 2502
|||||
Db 2238 CACACCTTTTGCAGAGAGAGCCCAAAAAGGAGCAGCAACAGCTTTGCTTTGTTGCTAGC 2297
|||||
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACTTTTAGATTTAGTTTC 2562
|||||

Db 2298 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2357
|||||
Qy 2563 AGCGATGATCTACGCGCTCTTCTGACAGAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
|||||
Db 2358 AGCGATGATCTACGCGCTCTTCTGACAGAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2417
|||||
Qy 2623 TTACAAGCCTCAAGTGTCTCAATGGTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTC 2682
|||||
Db 2418 TTACAAGCCTCAAGTGTCTCAATGGTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTC 2477
|||||
Qy 2683 TTCAAGTCTCATCTGCCCATCAAGCTTCTTGCAGCCAGCAGTCTTTCACAACTTATCTGG 2742
|||||
Db 2478 TTCAAGTCTCATCTGCCCATCAAGCTTCTTGCAGCCAGCAGTCTTTCACAACTTATCTGG 2537
|||||
Qy 2743 GAGTTTTCAGAACTCTCTTCAAGTGTAGTTTCAAGTGAAGAGAGGGTCTTCCAGTTT 2802
|||||
Db 2538 GAGTTTTCAGAACTCTCTTCAAGTGTAGTTTCAAGTGAAGAGAGGGTCTTCCAGTTT 2597
|||||
Qy 2803 GGAGAAAAAGGAGGTTCCAGAGTAGATTTTAGCATAACTCAATTCCTAAGGAATCTTGG 2862
|||||
Db 2598 GGAGAAAAAGGAGGTTCCAGAGTAGATTTTAGCATAACTCAATTCCTAAGGAATCTTGG 2657
|||||
Qy 2863 ACTTGAGCACCTTAATGGATATATTTGAGAGAGAAACAGATCACTTTTGATGTATTAGTTGA 2922
|||||
Db 2658 ACTTGAGCACCTTAATGGATATATTTGAGAGAGAAACAGATCACTTTTGATGTATTAGTTGA 2717
|||||
Qy 2923 GATGGGGCAACAGGAGCTGAAGSAGATTGGAATCAATAGTCTTATGACATAGGCACAACT 2982
|||||
Db 2718 GATGGGGCAACAGGAGCTGAAGSAGATTGGAATCAATAGTCTTATGACATAGGCACAACT 2777
|||||
Qy 2983 AATTAAGAGTGCAGAGACTTATCTCCGAGCAACAGGCTTTAAACCATATATTAACTTT 3042
|||||
Db 2778 AATTAAGAGTGCAGAGACTTATCTCCGAGCAACAGGCTTTAAACCATATATTAACTTT 2837
|||||
Qy 3043 GAACACTCTGTTAGTGAACAAATCTTATAGATCTGTCTCTCTGATGATAAAGAGTTTCA 3102
|||||
Db 2838 GAACACTCTGTTAGTGAACAAATCTTATAGATCTGTCTCTCTGATGATAAAGAGTTTCA 2897
|||||
Qy 3103 GTCGTGGAGGAAGAGATGCAAGTACAGTTCGAGAGCAGAGATGGAGGTATGCGAGG 3162
|||||
Db 2898 GTCGTGGAGGAAGAGATGCAAGTACAGTTCGAGAGCAGAGATGGAGGTATGCGAGG 2957
|||||
Qy 3163 TGAATCTTCAACAGATACAAATATCTCAAGATTCAGAGGTTTGAAGAAACATATG 3222
|||||
Db 2958 TGAATCTTCAACAGATACAAATATCTCAAGATTCAGAGGTTTGAAGAAACATATG 3017
|||||
Qy 3223 GGAAGATACACTCACCGGAGAAAAAGGTTTCTGAAGAAAAACCAACCATGCAATGA 3282
|||||
Db 3018 GGAAGATACACTCACCGGAGAAAAAGGTTTCTGAAGAAAAACCAACCATGCAATGA 3077
|||||
Qy 3283 ACAGATGCTATTTTCATGGGTCTCTTTTGTGAATGCAATATCCACAAAGGCTTTGATGA 3342
|||||
Db 3078 ACAGATGCTATTTTCATGGGTCTCTCTTTGTGAATGCAATATCCACAAAGGCTTTGATGA 3137
|||||
Qy 3343 AAGCATGCTTACATAGTGGTGTATGTTTGGAGCTGGCATTTATTTTCTGCTGAAAACCTCTTC 3402
|||||
Db 3138 AAGCATGCTTACATAGTGGTGTATGTTTGGAGCTGGCATTTATTTTCTGCTGAAAACCTCTTC 3197
|||||
Qy 3403 CAAAAGCAATCAATATATGATATGGAATTTGGAGGAGTACTGGGGTGCAGTTTCAACAAGA 3462
|||||
Db 3198 CAAAAGCAATCAATATGATATGGAATTTGGAGGAGTACTGGGTGTCCAGTTTCAACAAGA 3257
|||||
Qy 3463 CAGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGGGGTAACTTTGGGAAAGTC 3522
|||||
Db 3258 CAGATCTTGTATACATTTGCCACAGGAGCTGCTCTTTTGGGGTAACTTTGGGAAAGTC 3317
|||||
Qy 3523 TTTCTCTCAGTTTCAAGTGAATGAAAATGGACATTTCTCTCCAGGTCACTACTCAGTCAAC 3582
|||||
Db 3318 TTTCTCTCAGTTTCAAGTGAATGAAAATGGACATTTCTCTCCAGGTCACTACTCAGTCAAC 3377
|||||
Qy 3583 TGGTAGGCCAGGTGAATGGCCCTAGCATTAGCTGAATATGTTTATTTACAGAGAGAAACA 3642
|||||
Db 3378 TGGTAGGCCAGGTGAATGGCCCTAGCATTAGCTGAATATGTTTATTTACAGAGAGAAACA 3437
|||||

QY 3643 GGCTTACCTCAGTATTTAATTACTTACCAGATTATGAGGCTGAAGGTATGTCGATGG 3702
 |||
 Db 3438 GGCTTACCTCAGTATTTAATTACTTACCAGATTATGAGGCTGAAGGTATGTCGATGG 3497
 |||
 QY 3703 ATAA 3706
 |||
 Db 3498 ATAA 3501
 |||
 RESULT 14
 AX029397
 LOCUS AX029397 3400 bp DNA linear PAT 16-SEP-2000
 DEFINITION Sequence 1 from Patent WO9915647.
 ACCESSION AX029397
 VERSION AX029397.1 GI:10190194
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3400)
 AUTHORS Sutherland, R.L. and Daly, R.J.
 TITLE A potential effector for the grb7 family of signalling proteins
 JOURNAL Patent: WO 9915647-A 1 01-APR-1999;
 GARVAN INST MED RES (AU) ; SUTHERLAND ROBERT LYNDAY (AU) ; DALY
 ROGER JOHN (AU)
 FEATURES
 source Location/Qualifiers
 1. .3400
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 1059 a 665 c 760 g 916 t
 ORIGIN
 Query Match 86.8%; Score 3297; DB 6; Length 3400;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 3303; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 482 ATTCCTCTTCATATGATGCTCTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTCGGA 541
 |||
 Db 1 ATTCCTCTTCATATGATGCTCTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTCGGA 60
 |||
 QY 542 CATGGTCGAGACCCCAATGCTCGAGATAAATGGAAATTAATCTCTCTCCATGAAGCTGCA 601
 |||
 Db 61 CATGGTCGAGACCCCAATGCTCGAGATAAATGGAAATTAATCTCTCTCCATGAAGCTGCA 120
 |||
 QY 602 ATTAAGGAAGAATGATGTTGGATGCTGTTACAGATGGAGCTGAGCCCAACCATC 661
 |||
 Db 121 ATTAAGGAAGAATGATGTTGGATGCTGTTACAGATGGAGCTGAGCCCAACCATC 180
 |||
 QY 662 CGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 721
 |||
 Db 181 CGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 240
 |||
 QY 722 ACTGGTGAATATAAGAAAGATGAATCTTTAGAAAGTGCAGAGTGGCAATGAAGAAAAA 781
 |||
 Db 241 ACTGGTGAATATAAGAAAGATGAATCTTTAGAAAGTGCAGAGTGGCAATGAAGAAAAA 300
 |||
 QY 782 ATGATGGCTCTACTCACACCATTAATGTCAACTGCCACGCAAGTATGGCAGAAAGTCA 841
 |||
 Db 301 ATGATGGCTCTACTCACACCATTAATGTCAACTGCCACGCAAGTATGGCAGAAAGTCA 360
 |||
 QY 842 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGAGTTGTACAGCTGTTACTGCAA 901
 |||
 Db 361 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGAGTTGTACAGCTGTTACTGCAA 420
 |||
 QY 902 CATGGAGCTGATGTCATGCTAAAGATAAAGGTGATCTGTTACCATTACACAATGCGCTGT 961
 |||
 Db 421 CATGGAGCTGATGTCATGCTAAAGATAAAGGTGATCTGTTACCATTACACAATGCGCTGT 480
 |||
 QY 962 TCTTATGGTCATATTGAAGTAACTGAACCTTTTGGTCAAGCATGGTGCCTGTGTAATGCA 1021
 |||
 Db 481 TCTTATGGTCATATTGAAGTAACTGAACCTTTTGGTCAAGCATGGTGCCTGTGTAATGCA 540
 |||

QY 1022 ATGGACTTGTGGCAATTTCACTCCTCTTCAATGAGGAGCTTCTTAAGAAACAGGCTTGAAGTA 1081
 |||
 Db 541 ATGGACTTGTGGCAATTTCACTCCTCTTCAATGAGGAGCTTCTTAAGAAACAGGCTTGAAGTA 600
 |||
 QY 1082 TGTTCCTCTCTTCTTAAGTTATGTCGAGACCCCAACACTCCTCAATTTGTACAAATAAAAGT 1141
 |||
 Db 601 TGTTCCTCTCTTCTTAAGTTATGTCGAGACCCCAACACTCCTCAATTTGAAGAATAAAAGT 660
 |||
 QY 1142 GCTATAGACTTGGCTCCACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAAGGC 1201
 |||
 Db 661 GCTATAGACTTGGCTCCACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAAGGC 720
 |||
 QY 1202 CACTCTCTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAACATCTCTCT 1261
 |||
 Db 721 CACTCTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAACATCTCTCT 780
 |||
 QY 1262 CTGGAATGCTGAATTTCAAGCATCTCAACACATGAACACAGATTCGATTTGTGCTGCT 1321
 |||
 Db 781 CTGGAATGCTGAATTTCAAGCATCTCAACACATGAACACAGATTCGATTTGTGCTGCT 840
 |||
 QY 1322 GCATCTCCATATCCCAAAAGAAAGCAAAATATGTGAATGTTGCTRAGAAAAGGACCAAC 1381
 |||
 Db 841 GCATCTCCATATCCCAAAAGAAAGCAAAATATGTGAATGTTGCTRAGAAAAGGACCAAC 900
 |||
 QY 1382 ATCAATGAAAAGACTAAAGAAATTTCTTGAATCTCTGACGCTGCGCATCTGAGAAAAGCTCAT 1441
 |||
 Db 901 ATCAATGAAAAGACTAAAGAAATTTCTTGAATCTCTGACGCTGCGCATCTGAGAAAAGCTCAT 960
 |||
 QY 1442 AATGATGTTGTTGAAGTAGTGGTGAACACATGAAGCAAGGTTAAAGTCTCTGGATTAATCTT 1501
 |||
 Db 961 AATGATGTTGTTGAAGTAGTGGTGAACACATGAAGCAAGGTTAAAGTCTCTGGATTAATCTT 1020
 |||
 QY 1502 GGTACAGATCTCTACACAGAGCTGCATATTTGGTTCATCTACAAACCTTGGCCCTACTC 1561
 |||
 Db 1021 GGTACAGATCTCTACACAGAGCTGCATATTTGGTTCATCTACAAACCTTGGCCCTACTC 1080
 |||
 QY 1562 CTGAGCTATGGGTGTGATCTTAACATTAATCCCTTTCAGGGCTTTACTGCTTTACAGATG 1621
 |||
 Db 1081 CTGAGCTATGGGTGTGATCTTAACATTAATCCCTTTCAGGGCTTTACTGCTTTACAGATG 1140
 |||
 QY 1622 GGAATGAAAATGTACAGCAACTCTCCCAAGAGGGTATCTCATAGGTATTTACAGAGGCA 1681
 |||
 Db 1141 GGAATGAAAATGTACAGCAACTCTCCCAAGAGGGTATCTCATAGGTATTTACAGAGGCA 1200
 |||
 QY 1682 GACAGCAATTTGCTGGAAGCTGCAAGGCTGAGAGTGTGAAACTGTGAAACTGTAAACAACTGTGT 1741
 |||
 Db 1201 GACAGCAATTTGCTGGAAGCTGCAAGGCTGAGAGTGTGAAACTGTGAAACTGTAAACAACTGTGT 1260
 |||
 QY 1742 ACTGTTACAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACCACCTTCATTTT 1801
 |||
 Db 1261 ACTGTTACAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACCACCTTCATTTT 1320
 |||
 QY 1802 GCAGCTGGGTATTAACAGAGTGTCCGTTGGTGAATATCTGCTACAGATGGAGCTGATGTG 1861
 |||
 Db 1321 GCAGCTGGGTATTAACAGAGTGTCCGTTGGTGAATATCTGCTACAGATGGAGCTGATGTG 1380
 |||
 QY 1862 CATGCTAAAGATAAAGAGGCTTGTACCTTTGCACAATGCATGTTCTTTATGGACATTAAT 1921
 |||
 Db 1381 CATGCTAAAGATAAAGAGGCTTGTACCTTTGCACAATGCATGTTCTTTACGGACATTAAT 1440
 |||
 QY 1922 GAAGTTGCAGAACTTCTTTTAAACATGGAGCAGTAGTAAATAGTCTGATTTTATGGAAA 1981
 |||
 Db 1441 GAAGTTGCAGAACTTCTTTTAAACATGGAGCAGTAGTAAATAGTCTGATTTATGGAAA 1500
 |||
 QY 1982 TTTACACCTTTTACATGAAGCAGCAGCAAGAAAATATGAAAATTTGCAAACTTCTGCTC 2041
 |||
 Db 1501 TTTACACCTTTTACATGAAGCAGCAGCAAGAAAATATGAAAATTTGCAAACTTCTGCTC 1560
 |||
 QY 2042 CAGCATGGTGCAGACCCCTACCAAAAACAGGATGGAATACTCCTTTGGATCTGTT 2101
 |||
 Db 1561 CAGCATGGTGCAGACCCCTACCAAAAACAGGATGGAATACTCCTTTGGATCTGTT 1620
 |||

QY	2102	AAAGATGGAGATACAGATATTCAGATCTGCTTAGGGGAGATGACGCTTTGCTAGATGCT	2161		Db	2701	AAATATTCAGATTCAGAAGGTTTGTAAACAAGAAACTATGGGAAAGATACACTCACCGG	2760	
Db	1621	AAAGATGGAGATACAGATATTCAGATCTGCTTAGGGGAGATGACGCTTTGCTAGATGCT	1680		QY	3242	AGAAAAGAGTTTCTGAAGAAAACCAACCATGCCAATGAACGAATGCTATTTTCATGGG	3301	
QY	2162	GCCAGAGAGGTTGTTTAGCCAGAGTGAAGAGTTGCTTCTCCTGATGAATGAATGTGC	2221		Db	2761	AGAAAAGAGTTTCTGAAGAAAACCAACCATGCCAATGAACGAATGCTATTTTCATGGG	2820	
Db	1681	GCCAGAGAGGTTGTTTAGCCAGAGTGAAGAGTTGCTTCTCCTGATGAATGAATGTGC	1740		QY	3302	TCTCCTTTTGTGAATGCAATTTATCCAAAAGGCTTTGATGAAGGCATGCGTACATAGT	3361	
QY	2222	CSCGATACCCAGGAGACATTCACACCTTTACATTTTAGCAGCTGGTTATATTAATTTA	2281		Db	2821	TCTCCTTTTGTGAATGCAATTTATCCAAAAGGCTTTGATGAAGGCATGCGTACATAGT	2880	
Db	1741	CSCGATACCCAGGAGACATTCACACCTTTACATTTTAGCAGCTGGTTATATTAATTTA	1800		QY	3362	GGTATGTTTGAGGCTGGCATTTATTTTGTCTGAAAACCTCTTCCAAAAGCAATCAATATGTA	3421	
QY	2282	GAAGTTGAGAGTATTTGTTTACAACACGAGAGCTGATGTAATGCCCAAGCAAGAGGAGA	2341		Db	2881	GGTATGTTTGAGGCTGGCATTTATTTTGTCTGAAAACCTCTTCCAAAAGCAATCAATATGTA	2940	
Db	1801	GAAGTTGAGAGTATTTGTTTACAACACGAGAGCTGATGTAATGCCCAAGCAAGAGGAGA	1860		QY	3422	TATGGAATTTGGAGGAGTACTGGGTGTCAGGTTCCAGTTCCAAAAGACAGATCTTGTGTACATTTGC	3481	
QY	2342	CTTATTCCTTTTACATAATGCCAGCATCTTACGGGCAATGTAGATGTAGCAGCTCTACTAATA	2401		Db	2941	TATGGAATTTGGAGGAGTACTGGGTGTCAGGTTCCAGTTCCAAAAGACAGATCTTGTGTACATTTGC	3000	
Db	1861	CTTATTCCTTTTACATAATGCCAGCATCTTACGGGCAATGTAGATGTAGCAGCTCTACTAATA	1920		QY	3482	CACAGGACGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGCTTTCTTCCGTCAGTTTCAGTGCA	3541	
QY	2402	AGATATAATGATGTCTCAATGCCAGGACAAATGGGCTTTTCACACCTTTTGCCACGAAGCA	2461		Db	3001	CACAGGACGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGCTTTCTTCCGTCAGTTTCAGTGCA	3060	
Db	1921	AGATATAATGATGTCTCAATGCCAGGACAAATGGGCTTTTCACACCTTTTGCCACGAAGCA	1980		QY	3542	ATGAAAATGGGCATCTCTCCAGGTCTATCTCAGTCATCTCAGTCACCTGGTAGGCCAGTGAAT	3601	
QY	2462	GCCAAAAGGGAGCAACAGCTTTTGCTGTTTGTGTAGCCCATGGAGCTGACCCGACT	2521		Db	3061	ATGAAAATGGGCATCTCTCCAGGTCTATCTCAGTCATCTCAGTCACCTGGTAGGCCAGTGAAT	3120	
Db	1981	GCCAAAAGGGAGCAACAGCTTTTGCTGTTTGTGTAGCCCATGGAGCTGACCCGACT	2040		QY	3602	GGCCTTAGCATTTAGCTGAATATGTTTATTTACAGAGGAGAACAGGCTTATCTCAGTATTTTA	3661	
QY	2522	CTTAAAAATCAGGAAGGACAAACAGCTTTAGATTTTACGGGATGATGTACAGCGCT	2581		Db	3121	GGCCTTAGCATTTAGCTGAATATGTTTATTTACAGAGGAGAACAGGCTTATCTCAGTATTTTA	3180	
Db	2041	CTTAAAAATCAGGAAGGACAAACAGCTTTAGATTTTACGGGATGATGTACAGCGCT	2100		QY	3662	ATTACTTACGAGATTATGAGGCTTGAAGGTATGTCGATGGATAAATAGTTATTTTAAAGA	3721	
QY	2582	CTTCTGACAGCAGCATGCCCCATCTGCTCTGCCCTCTTGTACAAGCCTCAAGTGCTC	2641		Db	3181	ATTACTTACGAGATTATGAGGCTTGAAGGTATGTCGATGGATAAATAGTTATTTTAAAGA	3240	
Db	2101	CTTCTGACAGCAGCCATGCCCCATCTGCTCTGCCCTCTTGTACAAGCCTCAAGTGCTC	2160		QY	3722	AACTAATTCACCTGAACCTTAAATCATCAAGCAGCAGTGCGCTCTACGTTTTACTCCTT	3781	
QY	2642	AATGTTGTGAGAGAGCCAGGAGCCACTGCAGATGCTCTCTCTCAGGTCCATCTAGCCCCA	2701		Db	3241	AACTAATTCACCTGAACCTTAAATCATCAAGCAGCAGTGCGCTCTACGTTTTACTCCTT	3300	
Db	2161	AATGTTGTGAGAGAGCCAGGAGCCACTGCAGATGCTCTCTCTCAGGTCCATCTAGCCCCA	2220		QY	3782	TGCTGAAAAAAA 3794		
QY	2702	TCAAGCCTTTCTGACCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTTCAGAACTGCT	2761		Db	3301	TGCTGAAAAAAA 3313		
Db	2221	TCAAGCCTTTCTGACCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTTCAGAACTGCT	2280		RESULT 15				
QY	2762	TCAGTAGTTTCAAGTGAACAGAGGCTGCTTCCAGTTTTCAGAGAAAAGAGGTTCCA	2821		AX062142				
Db	2281	TCAGTAGTTTCAAGTGAACAGAGGCTGCTTCCAGTTTTCAGAGAAAAGAGGTTCCA	2340		LOCUS				
QY	2822	GGAGTAGATTTAGCATAACTCAATTCCTAAGGAATCTTGACACTTGACCACTAATGGAT	2881		DEFINITION				
Db	2341	GGAGTAGATTTAGCATAACTCAATTCGTAGGAATCTTGACACTTGACCACTAATGGAT	2400		Sequence 1 from Patent WO0100849.				
QY	2882	ATATTTGAGAGAGACAGATCACTTTGGATGATTAGTTGAGATGGGACAGAGGAGCTG	2941		AX062142				
Db	2401	ATATTTGAGAGAGACAGATCACTTTGGATGATTAGTTGAGATGGGACAGAGGAGCTG	2460		ACCESSION				
QY	2942	AGGAGATTGGAATCAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCCAGAGA	3001		AX062142.1				
Db	2461	AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCCAGAGA	2520		VERSION				
QY	3002	CTTATCTCCGGACAAACAGGCTCTTAACCCATATTTAACTTTGAACACCTCTGTTAGTGGA	3061		KEYWORDS				
Db	2521	CTTATCTCCGGACAAACAGGCTCTTAACCCATATTTAACTTTGAACACCTCTGTTAGTGGA	2580		human.				
QY	3062	ACAATTTCTTATAGATCTGCTCCTGATGATAAAGAGTTTTCAGTCTGTGGAGGAAGATG	3121		ORGANISM				
Db	2581	ACAATTTCTTATAGATCTGCTCCTGATGATAAAGAGTTTTCAGTCTGTGGAGGAAGATG	2640		Homo sapiens				
QY	3122	CAAAGTACAGTTTCGAGAGCAGACAGATGGAGTCTATGAGTGGATCTTCAACAGATAC	3181		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Db	2641	CAAAGTACAGTTTCGAGAGCAGACAGATGGAGTCTATGAGTGGATCTTCAACAGATAC	2700		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
QY	3182	AATATTTCTCAAGATTGAGAAGGTTTGTAAACAAGAACTATGGGAAAGATACACTCACCGG	3241		REFERENCE				
					1 (bases 1 to 3508)				
					Christenson, E., Demaggio, A.J., Goldman, P.S. and Mcelligott, D.L.				
					Tankyrase2 materials and methods				
					Patent: WO 0100849-A 1 04-JAN-2001;				
					ICOS CORPORATION (US)				
					Location/Qualifiers				
					1. 3508				
					/organism="Homo sapiens"				
					/db_xref="taxon:9606"				
					<2..>3508				
					/note="unnamed protein product"				
					/codon_start=1				
					/protein_id="CAC25052.1"				
					/db_xref="GI:12540046"				
					/translation="ARMSRRRCAGGGAACACAAAEVPAARELFEACRNGDVERVK				
					RLVTPKVNRSRDTRGRKSTPLHFAAGFGKDVYLLQNGANVQRDGLLPLHNAC				
					SFGHAEVYNLLRHGADPNARDNNWYTPLHEAAIKGKIDVCIQLQHAEPYIRITDG				
					RTALDLADPSAKAVLTGEYKKDELLSARSNGEEKMALLTPLNYNCHASDRKSTPL				

HLAAGYNRVIVQLLQHGADVHAKOKDVLPLHNAACSYGHVEVTELLVKGACVNA
 DMSFTPLHEAASNRVVECSLLISYKADPTLLNCHNKSAIDLAPTPOKLERLAEYFK
 GHSILOAREADVTIRIKKHLISLEWNEKHPQTHETALHCAASAPYKPKOICELLRK
 GANINKEKTELPPLHVASSEKAHNDVVEVVKHEAKVNALDNLGOTSUHLRAATCGHLQ
 TCRLLSTGSDPNIIISLQGTALOMNENVOQLLQEGISLGNSEADQRLLEAAKAGDV
 ETVKLLCTQSVNCRDLEGRSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGLVPL
 HNACSYGHEVALLVKGAVNVADLWKFPLHEAAKKGKYEICKLLOLHQAQDVH
 NRCDNTPLDLVKGDDTDIQLLRGDAALLDAKKGKLVKLLSPDNVNCRTDQGRK
 STPLHAAAGYNNLEAERYLLOHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNAC
 VNATDKWAPPLHEAAKGRTOICALLLAHGADPTLKNQEGPTPLDLVSADVYSALLT
 AAMPSPALPCYKQVNLGVRPCATADALUSSPSSLSAASSLNDLSGSELS
 VVSSSGTGASSLEKKEVPQDFSIQFVRLNGLHEHMDIFEREQITLDLVPMHKE
 LKETINAYGHRHLIKGVRLISGQQLNPLYTLNTSGSGTILDLSPDKDFQSV
 EEMOSTVREHRDGHAGGIFNRYNILKIQKCNKWLERYTHRRKEVSEENHNHNER
 MLPHGSPFVNAIHKGFDERHAYIGGMFGAGIYFAENS SKSNQYVVGIGGGTGPCVHK
 DRSCYICHRQLLFCRVTLKSFQFSAMKMAHSPPHSHSVTGRPSVNGLALEYIYR
 GEQAYPELTIYQIMRPEGMDVQ*

BASE COUNT 1046 a 710 c 843 g 909 t
 ORIGIN

Query Match 86.8%; Score 3296.2; DB 6; Length 3508;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3298; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	403	AGGTTTGGCGGAAGACGTAGTTGAATATTGCTTCAGAAATGGTCAAGTGTCCAAAGC	462
Db	208	AGGTTTGGCGGAAGACGTAGTTGAATATTGCTTCAGAAATGGTCAAGTGTCCAAAGC	267
Qy	463	AGTGATGATGGGGCCCTTATTCCTTCATTAATGATGCTCTTTTGGTCATGCTGAAGT	522
Db	268	AGTGATGATGGGGCCCTTATTCCTTCATTAATGATGCTCTTTTGGTCATGCTGAAGT	327
Qy	523	AGTCAATCTCTTTTGGCAGATGGTGACAGCCCAATCCTCGAGATAATTCGAATTATAC	582
Db	328	AGTCAATCTCTTTTGGCAGATGGTGACAGCCCAATCCTCGAGATAATTCGAATTATAC	387
Qy	583	TCCTCTCCATGAAGTGCATTAAGAGAAAGATTGATGTTTGCATTTGCTGTTTACAGCA	642
Db	388	TCCTCTCCATGAAGTGCATTAAGAGAAAGATTGATGTTTGCATTTGCTGTTTACAGCA	447
Qy	643	TGGAGCTGAGCCAAACATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC	702
Db	448	TGGAGCTGAGCCAAACATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC	507
Qy	703	ATCTGCGCAAGCAGTCTACTGTTGTAATAAGAAAGATGAATCTTAGAAAGTGCCAG	762
Db	508	ATCTGCGCAAGCAGTCTACTGTTGTAATAAGAAAGATGAATCTTAGAAAGTGCCAG	567
Qy	763	GAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTAATGTCACCTGCCAGCG	822
Db	568	GAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTAATGTCACCTGCCAGCG	627
Qy	823	AGTGATGGCAGAAAGTCACTCCATTTACATTTGGCAGCAGGATATAACAGATAAAGAT	882
Db	628	AGTGATGGCAGAAAGTCACTCCATTTACATTTGGCAGCAGGATATAACAGATAAAGAT	687
Qy	883	TGTACAGCTGTTTACTGCAACATGGAGCTGATGCTCATGCTAAAGATAAAGGTGATCTGGT	942
Db	688	TGTACAGCTGTTTACTGCAACATGGAGCTGATGCTCATGCTAAAGATAAAGGTGATCTGGT	747
Qy	943	ACCATTACCAATGGCTGTCTTATGGTCATTTAAGTAACTGAACCTTTTGGTCAAGCA	1002
Db	748	ACCATTACCAATGGCTGTCTTATGGTCATTTAAGTAACTGAACCTTTTGGTCAAGCA	807
Qy	1003	TGGTGGCTGTGAATGAATGGACTTTGGCAATTTCACTCTTTCATGAGGAGCTTC	1062
Db	808	TGGTGGCTGTGAATGAATGGACTTTGGCAATTTCACTCTTTCATGAGGAGCTTC	867
Qy	1063	TAAAGAACAGGTTGAAGTATGTTCTCTTCTTAAGTTATGGTCAGACCCCAACACTGCT	1122
Db	868	TAAAGAACAGGTTGAAGTATGTTCTCTTCTTAAGTTATGGTCAGACCCCAACACTGCT	927
Qy	1123	CAATTGTCAATAAAGTGTATAGACTTTGGCTCCCAACACAGTTTAAAGAAAGATT	1182

Db	928	CAATTGTGCACATAAAGTGCTATAGACTTGGCTCCACACACAGTATAAAGAAAAGATT	987
Qy	1183	AGCATATGAATTTAAAGGCCACATCGTTGCTGCAAGCTGCACGAGAGCTGATGTTACTCG	1242
Db	988	AGCATATGAATTTAAAGGCCACATCGTTGCTGCAAGCTGCACGAGAGCTGATGTTACTCG	1047
Qy	1243	AATCAAAAACATCTCTCTCTGGAATTTCAAGCATCTCTCAACACATGAAC	1302
Db	1048	AATCAAAAACATCTCTCTCTGGAATTTCAAGCATCTCTCAACACATGAAC	1107
Qy	1303	AGCATTCATTTGCTGCTGCATCTCCATATCCAAAGAAAGCAATATGTGAATGTT	1362
Db	1108	AGCATTCATTTGCTGCTGCATCTCCATATCCAAAGAAAGCAATATGTGAATGTT	1167
Qy	1363	GCTAAGAAAGAGGACCAACATCAATGAAAGACATAAGAAATTTCTTGACTCTCTGACGCT	1422
Db	1168	GCTAAGAAAGAGGACCAACATCAATGAAAGACATAAGAAATTTCTTGACTCTCTGACGCT	1227
Qy	1423	GGCATCTGAGAAAGCTCATATGATGTTTGAAGTAGTGGTGAACATGAAGCAAGGT	1482
Db	1228	GGCATCTGAGAAAGCTCATATGATGTTTGAAGTAGTGGTGAACATGAAGCAAGGT	1287
Qy	1483	TAATGCTCTGGATAAATCTTGGTCAGACTTCTCTACAGAGCTGCATATTTGTTGTCATCT	1542
Db	1288	TAATGCTCTGGATAAATCTTGGTCAGACTTCTCTACAGAGCTGCATATTTGTTGTCATCT	1347
Qy	1543	ACAAACCTGCGGCTACTCTGAGCTATGGTGTGATCTTAAACATATATCCCTTCAGGG	1602
Db	1348	ACAAACCTGCGGCTACTCTGAGCTATGGTGTGATCTTAAACATATATCCCTTCAGGG	1407
Qy	1603	CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCCAGAGGCTATCTC	1662
Db	1408	CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCCAGAGGCTATCTC	1467
Qy	1663	ATTAGGTAATTCAGAGCAGACAGACAATTTGCTGGAAGCTGCAAGAGCTGGAGATGCTGA	1722
Db	1468	ATTAGGTAATTCAGAGCAGACAGACAATTTGCTGGAAGCTGCAAGAGCTGGAGATGCTGA	1527
Qy	1723	AACGTAAAAAAACTGTGTACTGTTACAGTGTCAACTGCAGAGACATTTGAAGGGCTCA	1782
Db	1528	AACGTAAAAAAACTGTGTACTGTTACAGTGTCAACTGCAGAGACATTTGAAGGGCTCA	1587
Qy	1783	GTCTACACCACTTCATTTGCGAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCT	1842
Db	1588	GTCTACACCACTTCATTTGCGAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCT	1647
Qy	1843	ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCGCTTGTACCTTTTGCACATGC	1902
Db	1648	ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCGCTTGTACCTTTTGCACATGC	1707
Qy	1903	ATGTTCTTATGGACATTTAAGCTTGCAGACTTCTTTGTTAAACATGAGCAGTAGTTAA	1962
Db	1708	ATGTTCTTATGGACATTTAAGCTTGCAGACTTCTTTGTTAAACATGAGCAGTAGTTAA	1767
Qy	1963	TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGATAATGA	2022
Db	1768	TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGATAATGA	1827
Qy	2023	AATTTGCAAACTTCTGCTCCAGCATGTCAGAGCCCTTACCAAAAAAAGAGGATGGA	2082
Db	1828	AATTTGCAAACTTCTGCTCCAGCATGTCAGAGCCCTTACCAAAAAAAGAGGATGGA	1887
Qy	2083	TACTCTCTTGGATCTCTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTGGGAGA	2142
Db	1888	TACTCTCTTGGATCTCTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTGGGAGA	1947
Qy	2143	TGACGCTTTGCTAGATGCTCCAAAGAGGTTGTTTACCCAGAGTGAAGATGTTCTTC	2202
Db	1948	TGACGCTTTGCTAGATGCTCCAAAGAGGTTGTTTACCCAGAGTGAAGATGTTCTTC	2007
Qy	2203	TCCTGATAATGATAAATGGCGGATACCCAGGAGACATTTCAACACTTTTACATTTAGC	2262
Db		TCCTGATAATGATAAATGGCGGATACCCAGGAGACATTTCAACACTTTTACATTTAGC	

Db 2008 TCCTGATAATGTAAATGCCCGGATACCCAAAGGCGAGACATCAACACACCTTTACATTTAGC 2067
QY 2263 AGCTGGTTATAATAATTTAGAAAGTTGCAGAGTATTTGTTTACAACACGAGCTGATGTGAA 2322
Db 2068 AGCTGGTTATAATAATTTAGAAAGTTGCAGAGTATTTGTTTACAACACGAGCTGATGTGAA 2127
QY 2323 TGCCCAAGACAAAGAGGAGCTTATTTCCCTTTACATTAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2128 TGCCCAAGACAAAGAGGAGCTTATTTCCCTTTACATTAATGCAGCATCTTACGGGCATGTAGA 2187
QY 2393 TGTACAGCTCTACTAATAAAGATATAATGCATGTCTCAATGCCACGAGCAAAATGGCTTT 2442
Db 2188 TGTACAGCTCTACTAATAAAGATATAATGCATGTCTCAATGCCACGAGCAAAATGGCTTT 2247
QY 2443 CACACCTTTGACGAAGACAGCCCAAAAGGAGCAACACAGCTTTGTGCTTTGCTTACG 2502
Db 2248 CACACCTTTGACGAAGACAGCCCAAAAGGAGCAACACAGCTTTGTGCTTTGCTTACG 2307
QY 2503 CCATGGAGCTGACCCGACTCTTTAAAATCAGGAAGCAAAACACCTTTAGATTTAGTTTC 2562
Db 2308 CCATGGAGCTGACCCGACTCTTTAAAATCAGGAAGCAAAACACCTTTAGATTTAGTTTC 2367
QY 2563 AGCGGATGATGTCAGGCTCTCTCACAGCAGCCATGCCCATCTGCTCTGCCCTCTTG 2622
Db 2368 AGCAGATGATGTCAGGCTCTCTCACAGCAGCCATGCCCATCTGCTCTGCCCTCTTG 2427
QY 2623 TTACAAGCTCAAGTCTCAATGGTGTGAGAAAGCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2428 TTACAAGCTCAAGTCTCAATGGTGTGAGAAAGCCAGGAGCCACTGCAGATGCTCTCTC 2487
QY 2683 TTCAGGTCATCTAGCCATCAAGCTTTCTCGACGCGAGCAGCTTTGACAACTTATCTG 2742
Db 2488 TTCAGGTCATCTAGCCATCAAGCTTTCTCGACGCGAGCAGCTTTGACAACTTATCTG 2547
QY 2743 GAGTTTTCAGAACTGCTCTTCAAGTCTAGTTCAAGTGAAGCAAGAGGCTTCCAGTTT 2802
Db 2548 GAGTTTTCAGAACTGCTCTTCAAGTCTAGTTCAAGTGAAGCAAGAGGCTTCCAGTTT 2607
QY 2803 GGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2608 GCAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2667
QY 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAAACAGATCATTGGATGTATTAGTTGA 2922
Db 2668 ACTTGAGCACCTAATGGATATATTTGAGAGAGAAACAGATCATTGGATGTATTAGTTGA 2727
QY 2923 GATGGGCAACAAGGAGCTGAAGGAGATGGAAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2728 GATGGGCAACAAGGAGCTGAAGGAGATGGAAATCAATGCTTATGGACATAGGCACAAACT 2787
QY 2983 AATTAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGCTTTAACCCATATTTAACTTT 3042
Db 2788 AATTAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGCTTTAACCCATATTTAACTTT 2847
QY 3043 GAACACCTCTGTTAGTGGAACTTCTTATAGATCTGCTCTCTGATGATAAAGAGTTTCA 3102
Db 2848 GAACACCTCTGTTAGTGGAACTTCTTATAGATCTGCTCTCTGATGATAAAGAGTTTCA 2907
QY 3103 GTCTGTGGAGGAGAGATGCAAAAGTACAGTTCGAGAGCAGAGATGGAGGTCATGCAGG 3162
Db 2908 GTCTGTGGAGGAGAGATGCAAAAGTACAGTTCGAGAGCAGAGATGGAGGTCATGCAGG 2967
QY 3163 TGGAACTCTTCAACAGATACAATATTTCTCAAGATTCAGAAAGTTTGTAAACAAGAACTATG 3222
Db 2968 TGGAACTCTTCAACAGATACAATATTTCTCAAGATTCAGAAAGTTTGTAAACAAGAACTATG 3027
QY 3223 GGAAGATACACTCACCGAGAGAAAGAACTTTCTGAAGAAAACCAACCATGCCAATGA 3282
Db 3028 GGAAGATACACTCACCGAGAGAAAGAACTTTCTGAAGAAAACCAACCATGCCAATGA 3087
QY 3283 ACGAATGCTTATTTTCATGGGCTCTCTTCTGTAATGCAATTTATCCACAAAGGCTTTGATGA 3342
Db 3088 ACGAATGCTTATTTTCATGGGCTCTCTTCTGTAATGCAATTTATCCACAAAGGCTTTGATGA 3147

QY 3343 AAGGCATCGGTACATAGGTGGTGTATGTTTGGAGCTGGCATTTATTTTGTGTAACACTCTTC 3402
Db 3148 AAGGCATCGGTACATAGGTGGTGTATGTTTGGAGCTGGCATTTATTTTGTGTAACACTCTTC 3207
QY 3403 CAAAGCAATCAATATGTATATGGAATTTGGAGAGGTACTGGGTGTCCAGTTTCACAAGA 3462
Db 3208 CAAAGCAATCAATATGTATATGGAATTTGGAGAGGTACTGGGTGTCCAGTTTCACAAGA 3267
QY 3463 CAGATCTTTGTTACATTTGCCACAGCAGTCTCTTTTCCCGGTAACCTTGGGAAAGTC 3522
Db 3268 CAGATCTTTGTTACATTTGCCACAGCAGTCTCTTTTCCCGGTAACCTTGGGAAAGTC 3327
QY 3523 TTTCTGTGAGTTTCAAGTCAATGAAAATGGCAATTTCTCTCCAGGTCACTCAGTCTAC 3582
Db 3328 TTTCTGTGAGTTTCAAGTCAATGAAAATGGCAATTTCTCTCCAGGTCACTCAGTCTAC 3387
QY 3583 TGGTAGGCCAGTGAATGGCTAGCTTACCTGAATATGTTTATACAGAGGAGAACA 3642
Db 3388 TGGTAGGCCAGTGAATGGCTAGCTTACCTGAATATGTTTATACAGAGGAGAACA 3447
QY 3643 GGCTTATCTCTGAGTATTTAAATTTACTTTACAGATTTATGAGGCCTGAAGGTATGTCGATGG 3702
Db 3448 GGCTTATCTCTGAGTATTTAAATTTACTTTACAGATTTATGAGGCCTGAAGGTATGTCGATGG 3507
QY 3703 A 3703
Db 3508 A 3508

Search completed: February 11, 2003, 18:25:58
Job time : 9476.36 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 12:17:23 ; Search time 717.206 Seconds
(without alignments)
11922.436 Million cell updates/sec

Title: US-09-843-159B-1

Perfect score: 3797

Sequence: 1 cttgaagacactggtttc.....ccttctgataaaaaaaa 3797

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3797	100.0	3797	22	AAA91487
2	3393.4	89.4	3816	22	AAA91488
3	3388.8	89.2	4275	21	AAC66825
4	3388.8	89.2	6019	22	AAC85294
5	3387.2	89.2	4127	22	AAF63926
6	3387.2	89.2	4406	22	AAF63930
7	3387.2	89.2	4992	22	AAF63953
8	3387.2	89.2	5002	22	AAF63952
9	3382.8	89.1	4512	22	AAD02578

10	3297	86.8	3400	20	AA253366
11	3296.2	86.8	3508	22	AAF63837
12	2969.6	78.2	4493	21	AAC66823
13	2968.2	78.2	4296	21	AAC66824
14	2774.6	73.1	5005	22	AAC89706
15	2439.8	64.3	2971	22	AAF63919
16	2439.8	64.3	3253	22	AAF63917
17	2357.8	62.1	3799	22	AAF63918
18	1813.4	47.8	3984	22	AAF63838
19	1813.4	47.8	4134	21	AAC66826
20	1813.4	47.8	4134	21	AAC66827
21	1813.4	47.8	4134	24	ABK48629
22	1794.8	47.3	3254	22	AAC85295
23	1449.6	38.2	4491	21	AAZ29632
24	1449.6	38.2	4491	24	ABK48634
25	1359.4	35.8	1568	22	AAF63925
26	1297.8	34.2	4657	21	AAZ29633
27	1297.8	34.2	4657	24	ABK48635
28	1240.4	32.7	1466	22	ABA08564
29	1172.2	30.9	1242	22	AAC89701
30	1168.4	30.8	3308	22	AAF63990
31	1084.8	28.6	1103	22	AAF63973
32	1076.8	28.4	3006	22	AAH16624
33	1008.8	26.6	5482	22	AAF63955
34	1008.8	26.6	5484	23	ABL08155
35	994	26.2	2409	22	AAK99406
36	990	26.1	1007	22	AAC89703
37	786	20.7	1344	21	AAC77578
38	784.2	20.7	838	22	AAH08231
39	748.6	19.7	10194	23	ABL08154
40	574	15.1	582	22	AAF63875
41	560	14.7	2793	21	AAF63928
42	518.4	13.7	520	21	AAC66828
43	459	12.1	460	22	AAH35131
44	459	12.1	460	22	AAF63839
45	396.8	10.5	402	22	AAH42791

ALIGNMENTS

RESULT 1
AAA91487

ID AAA91487 standard; DNA; 3797 BP.

XX AAA91487;

AC AAA91487;

DT 06-AUG-2001 (first entry)

XX Tankyrase homologue isotype 1 coding sequence.

XX Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine; ds.

XX Unidentified.

OS Unidentified.

XX Key Location/Qualifiers

FT CDS 404..3706

FT /*tag= a

FT /partial

FT /product= "TaHo1"

FT /note= "Tankyrase homologue isotype 1; No start codon given"

FT FT

XX WO200130987-A2.

XX PD 03-MAY-2001.

XX PF 25-OCT-2000; 2000WO-US41528.

XX PR 25-OCT-1999; 99US-0427154.

PA	(RIGE-) RIGEL PHARM INC.	
XX	Luo Y, Chan E, Xu X, Huang B;	
PI	WPI: 2001-300503/31.	
XX	P-PSDB: AAY97748.	
DR	Novel recombinant cell cycle polypeptide, tankyrase H useful for	
PT	inducing or preventing cell proliferation in cells, and for diagnosing,	
PT	treating or preventing cell cycle associated disorders such as cancer	
PT	.	
XX	Claim 13; Fig 1; 63pp; English.	
PS	This sequence encodes the Tankyrase homologue isotype 1 (TaHo-1) protein	
XX	of the invention. The invention also relates to the TaHo-2 protein.	
CC	The TaHo proteins are useful for inducing or preventing cell	
CC	proliferation in cells, and in the study or treatment of conditions	
CC	mediated by the cell cycle proteins, such as to diagnose, treat or	
CC	prevent cell cycle associated disorders, preferably cancer. The TaHo	
CC	coding sequences are useful as hybridisation probes, in chromosome and	
CC	gene mapping and in the generation of anti-sense DNA and RNA. The coding	
CC	sequences are also useful for the preparation of TaHo, for generating	
CC	either transgenic animals or knock out animals which, in turn, are useful	
CC	in a development and screening of therapeutically useful agents, in gene	
CC	therapy, as vaccine, and for construction of hybridisation probes for	
CC	mapping the gene which encodes TaHo and for the genetic analysis of	
CC	individuals with genetic disorders. The TaHo proteins, and their coding	
XX	sequences are useful in screening assays.	
XX	Sequence 3797 BP; 1153 A; 732 C; 850 G; 1062 T; 0 other;	
QY	Query Match	100.0%; Score 3797; DB 22; Length 3797;
DB	Best Local Similarity	100.0%; Pred. No. 0;
QY	Matches 3797; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CTTTGAAGACACTGGATTTCATACATTTTGGCTGGGGTTATCTCTCTGTGTCCTACTACAT	60
DB	1 CTTTGAAGACACTGGATTTCATACATTTTGGCTGGGGTTATCTCTCTGTGTCCTACTACAT	60
QY	61 AGACAAATATTAGCTGTGAGCAGATCTTTTGGCTGGGGTTATCTCTCTGTGTCCTACTACAT	120
DB	61 AGACAAATATTAGCTGTGAGCAGATCTTTTGGCTGGGGTTATCTCTCTGTGTCCTACTACAT	120
QY	121 CAGAAACATCTGAGATAGATGCGGAAAGCAATTTCTAGCAAGAGTTTGTGCACGTGA	180
DB	121 CAGAAACATCTGAGATAGATGCGGAAAGCAATTTCTAGCAAGAGTTTGTGCACGTGA	180
QY	181 TCATAAGGTTGCTGATTACATPATTAAAGTTTATCTTTGAACATCTGAAATGTATACA	240
DB	181 TCATAAGGTTGCTGATTACATPATTAAAGTTTATCTTTGAACATCTGAAATGTATACA	240
QY	241 TACTAAATATCAGAACTCTATTGTAGAGTCAGAAACATTTGAACCTTTGAGCTTTCAGTC	300
DB	241 TACTAAATATCAGAACTCTATTGTAGAGTCAGAAACATTTGAACCTTTGAGCTTTCAGTC	300
QY	301 ACTATTATTGATTCTTTCTTTGAGTTAGCAGTAGTACCACCCCAAGCACTGCTTAGGT	360
DB	301 ACTATTATTGATTCTTTCTTTGAGTTAGCAGTAGTACCACCCCAAGCACTGCTTAGGT	360
QY	361 ACCACTGCTGCTAGTGAGAGTCCTCTGGCTTTATCAATTAAGTTTGGCGGAAAGA	420
DB	361 ACCACTGCTGCTAGTGAGAGTCCTCTGGCTTTATCAATTAAGTTTGGCGGAAAGA	420
QY	421 CGTAGTTGAATATTGCTTCAGAAATGGTGAAGTGTGCAAGCACTGATGATGGGGCT	480
DB	421 CGTAGTTGAATATTGCTTCAGAAATGGTGAAGTGTGCAAGCACTGATGATGGGGCT	480
QY	481 TATTCTCTTCAATATGATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCTTTGGG	540
DB	481 TATTCTCTTCAATATGATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCTTTGGG	540
QY	541 ACATGGTCAGACCCCAATGCTCGAGATAATTTGGAATATATCTCTCTCCATGAAGCTGC	600

Db 1621 GGGAAATGAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGC 1680
QY 1681 AGACAGACAAATGCTGGAGCTGCAAGGCTGGAGATCTCGAAACTGTFAAAAAACACTGTG 1740
Db 1681 AGACAGACAAATGCTGGAGCTGCAAGGCTGGAGATCTCGAAACTGTFAAAAAACACTGTG 1740
QY 1741 TACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACCACATTCATTT 1800
Db 1741 TACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACCACATTCATTT 1800
QY 1801 TGCAGCTGGGTATTAACAGAGTGTCCGTTGGTGAATATCTGCTACAGCATGGAGCTGATGT 1860
Db 1801 TGCAGCTGGGTATTAACAGAGTGTCCGTTGGTGAATATCTGCTACAGCATGGAGCTGATGT 1860
QY 1861 GCATGCTAAAGATAAAGAGGCGCTTGTACCTTTGCACAAATGCATGTTCTTTATGGACATTA 1920
Db 1861 GCATGCTAAAGATAAAGAGGCGCTTGTACCTTTGCACAAATGCATGTTCTTTATGGACATTA 1920
QY 1921 TGAAGTTGCACAACCTCTCTGTTAAACATGGAGCAGTAGTTAAAGTAGCTGATTTATGGAA 1980
Db 1921 TGAAGTTGCACAACCTCTCTGTTAAACATGGAGCAGTAGTTAAAGTAGCTGATTTATGGAA 1980
QY 1981 ATTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCGCT 2040
Db 1981 ATTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCGCT 2040
QY 2041 CCAGCATGGTGCAGACCGCTACCAAAAAACAGGGATGGAATACTCCTTTGGATCTTGT 2100
Db 2041 CCAGCATGGTGCAGACCGCTACCAAAAAACAGGGATGGAATACTCCTTTGGATCTTGT 2100
QY 2101 TAAAGATGGAGATACATATTCAGATCTGCTTTAGCCAGTGAAGTACTCCTTTGGATCTTGT 2160
Db 2101 TAAAGATGGAGATACATATTCAGATCTGCTTTAGCCAGTGAAGTACTCCTTTGGATCTTGT 2160
QY 2161 TGCCAAAGAGGGTTGTTAGCCAGTGAAGAGTGTCTCTCTCATATTAATTAATTTG 2220
Db 2161 TGCCAAAGAGGGTTGTTAGCCAGTGAAGAGTGTCTCTCTCATATTAATTAATTTG 2220
QY 2221 CCGCGATTACCAAGGAGACATTCACACCTTTACATTTAGCAGCTGGTTATTAATTAATTT 2280
Db 2221 CCGCGATTACCAAGGAGACATTCACACCTTTACATTTAGCAGCTGGTTATTAATTAATTT 2280
QY 2281 AGAAGTTGCAGAGTATTTGTTACACACGGAGCTGATGCAATGCCCAACAGCAAGAGG 2340
Db 2281 AGAAGTTGCAGAGTATTTGTTACACACGGAGCTGATGCAATGCCCAACAGCAAGAGG 2340
QY 2341 ACTTATTCCTTTTACATAATGACGATCTTACGGGCATGTAGATGTAGCAGCTCTACTAAT 2400
Db 2341 ACTTATTCCTTTTACATAATGACGATCTTACGGGCATGTAGATGTAGCAGCTCTACTAAT 2400
QY 2401 AAAGTATAATGCAATGCTCAATGCCAGCACAATAAGGGCTTTTCACACCTTTGCACGAAGC 2460
Db 2401 AAAGTATAATGCAATGCTCAATGCCAGCACAATAAGGGCTTTTCACACCTTTGCACGAAGC 2460
QY 2461 AGCCCAAAAGGGACGACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGAC 2520
Db 2461 AGCCCAAAAGGGACGACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGAC 2520
QY 2521 TCTTAAAAATCAGGAGGACAACACCTTTAGATTTAGTTTTCAGCGGATGATGTACGCGC 2580
Db 2521 TCTTAAAAATCAGGAGGACAACACCTTTAGATTTAGTTTTCAGCGGATGATGTACGCGC 2580
QY 2581 TCTTCTTGACAGCAGCATGCCCCCATCTGCTGCTGCTGCTTCTGTTACAGCCCTCAAGTGT 2640
Db 2581 TCTTCTTGACAGCAGCATGCCCCCATCTGCTGCTGCTGCTTCTGTTACAGCCCTCAAGTGT 2640
QY 2641 CAATGCTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTTTCAGGTGCCATCTAGCCC 2700
Db 2641 CAATGCTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTTTCAGGTGCCATCTAGCCC 2700
QY 2701 ATCAAGCCTTTCTGAGCCAGCAGTCTTGACAACTTATCTGGAGTGTTCAGAACTGTC 2760
Db 2701 ATCAAGCCTTTCTGAGCCAGCAGTCTTGACAACTTATCTGGAGTGTTCAGAACTGTC 2760

QY 2761 TTCAGTAGTTAGTTCAAGTGAACAGAGGGTGTCTCCAGTTTCGAGAAAAAGAGGTTC 2820
Db 2761 TTCAGTAGTTAGTTCAAGTGAACAGAGGGTGTCTCCAGTTTCGAGAAAAAGAGGTTC 2820
QY 2821 AGGAGTAGATTTTAGCATAACTCAATTCGTTAAGGAATCTTGGACTTGGACCCTAATGGA 2880
Db 2821 AGGAGTAGATTTTAGCATAACTCAATTCGTTAAGGAATCTTGGACTTGGACCCTAATGGA 2880
QY 2881 TATATTGAGAGAGAACAGATCACTTTTGGATGTATTAGTTGAGATGGGCGACAGGAGCT 2940
Db 2881 TATATTGAGAGAGAACAGATCACTTTTGGATGTATTAGTTGAGATGGGCGACAGGAGCT 2940
QY 2941 GAAGGAGATTGGAAATCAATGCTTTATGAGCATAGGCAACAACAAATTAAGGAGTGCAGAG 3000
Db 2941 GAAGGAGATTGGAAATCAATGCTTTATGGATATAGGCAACAACAAATTAAGGAGTGCAGAG 3000
QY 3001 ACTTATCTCCGGACAACAAGGCTTTAAACCATATTTAACTTTGAACACCTCTGGTAGTGG 3060
Db 3001 ACTTATCTCCGGACAACAAGGCTTTAAACCATATTTAACTTTGAACACCTCTGGTAGTGG 3060
QY 3061 AACAAATCTTATAGATCTGCTCCTCATATAAAGAGTTTTCAGTCTGTGGAGAAAGAT 3120
Db 3061 AACAAATCTTATAGATCTGCTCCTCATATAAAGAGTTTTCAGTCTGTGGAGAAAGAT 3120
QY 3121 GCAAGGTACAGTTCGAGAGCAGAGATGGAGTCTATGCAGTGGAAATCTTCAACAGATA 3180
Db 3121 GCAAGGTACAGTTCGAGAGCAGAGATGGAGTCTATGCAGTGGAAATCTTCAACAGATA 3180
QY 3181 CAATATCTCAAGTTTCAGAAGTGTGTAAACAAGAACTATGGGAAAGATACACCTCACCG 3240
Db 3181 CAATATCTCAAGTTTCAGAAGTGTGTAAACAAGAACTATGGGAAAGATACACCTCACCG 3240
QY 3241 GAGAAAAGAGTTTCTGAAAGAAAACCAACCATGCCAATGAACAAATGCTATTTTCATGG 3300
Db 3241 GAGAAAAGAGTTTCTGAAAGAAAACCAACCATGCCAATGAACAAATGCTATTTTCATGG 3300
QY 3301 GTCTCCTTTTGTGAATGCAATATCCACAAAGGCTTTGATGAAGGCGATGCGTACATAGG 3360
Db 3301 GTCTCCTTTTGTGAATGCAATATCCACAAAGGCTTTGATGAAGGCGATGCGTACATAGG 3360
QY 3361 TGGTATGTTTGGAGTGGCAATTTATTTTGTGAAAACTCTTCCAAAAGCAATCAATATGT 3420
Db 3361 TGGTATGTTTGGAGTGGCAATTTATTTTGTGAAAACTCTTCCAAAAGCAATCAATATGT 3420
QY 3421 ATATGSAATTTGAGCAGGACTGCTGGGTGCCAGTTCACAAAGACAGATCTTGTACATTTG 3480
Db 3421 ATATGSAATTTGAGCAGGACTGCTGGGTGCCAGTTCACAAAGACAGATCTTGTACATTTG 3480
QY 3481 CCACAGGAGCTGCTCTTTTGGCGGTAACCTTTGGGAAAGTCTTTCCCTGCAAGTTCAGTGC 3540
Db 3481 CCACAGGAGCTGCTCTTTTGGCGGTAACCTTTGGGAAAGTCTTTCCCTGCAAGTTCAGTGC 3540
QY 3541 AATGAAATGGCACAATCTCCTCCAGGTCATCACTGAGTGGCCAGCTGATAA 3600
Db 3541 AATGAAATGGCACAATCTCCTCCAGGTCATCACTGAGTGGCCAGCTGATAA 3600
QY 3601 TGGCTAGCATTTAGCTGAATATGTTTACAGAGGAGCAACAGGCTTATCTGAGTATTT 3660
Db 3601 TGGCTAGCATTTAGCTGAATATGTTTACAGAGGAGCAACAGGCTTATCTGAGTATTT 3660
QY 3661 AATTACTTACCAGATTATGAGGCGCTGAAGTATGCTGCATGATAAATAGTATTTTAAAG 3720
Db 3661 AATTACTTACCAGATTATGAGGCGCTGAAGTATGCTGCATGATAAATAGTATTTTAAAG 3720
QY 3721 AAACATAATTCACCTGAACCTAAATCATCAAGCAGAGTGGCTCTACGCTTTTACTCCT 3780
Db 3721 AAACATAATTCACCTGAACCTAAATCATCAAGCAGAGTGGCTCTACGCTTTTACTCCT 3780
QY 3781 TTGCTGAAAAAATAA 3797
Db 3781 TTGCTGAAAAAATAA 3797

Db 1562 ACAAACTGCGGCTACTCTCTAGCTATGGGTGTGATCTCAATATATATCCCTTCAGGG 1621
QY 1603 CTTTACTGCTTTACAGATGGGAATGAAATGACAGCAACTCTCTCAAGAGGGTATCTC 1662
Db 1622 CTTTACTGCTTTACAGATGGGAATGAAATGACAGCAACTCTCTCAAGAGGGTATCTC 1681
QY 1663 ATTAGTAAATTCAGAGCGACAGACAATTCGTGGAAGCTGCAAGAGCTGGAGATGTCA 1722
Db 1682 ATTAGTAAATTCAGAGCGACAGACAATTCGTGGAAGCTGCAAGAGCTGGAGATGTCA 1741
QY 1723 AACTGTAAAAAACTGTGTACTGTTACAGAGTGCACATTCGAGAGCAATTCGAGGGGTCA 1782
Db 1742 AACTGTAAAAAACTGTGTACTGTTACAGAGTGCACATTCGAGAGCAATTCGAGGGGTCA 1801
QY 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATACACAGAGTGTCCGTGGTGAATATCTGCT 1842
Db 1802 GTCTACACCACTTCATTTTGCAGCTGGGTATACACAGAGTGTCCGTGGTGAATATCTGCT 1861
QY 1843 ACAGCATGGAGCTGTGTGATGCTAAAGATAAAGAGAGCTTGTACCTTTGCACAATGC 1902
Db 1862 ACAGCATGGAGCTGTGTGATGCTAAAGATAAAGAGAGCTTGTACCTTTGCACAATGC 1921
QY 1903 ATGTCTCTTATGGACATATATCAAGTTCGACAACTTCTTGTAAACATTCGAGCAGTAGTTAA 1962
Db 1922 ATGTCTCTTATGGACATATATCAAGTTCGACAACTTCTTGTAAACATTCGAGCAGTAGTTAA 1981
QY 1963 TGTAGCTGATTTTATGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATCA 2022
Db 1982 TGTAGCTGATTTTATGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATCA 2041
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCACCAAAACAGGGATGGAA 2082
Db 2042 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCACCAAAACAGGGATGGAA 2101
QY 2083 TACTCTCTTGGATCTTTAAAGATGGAGATACAGATATTCAGATCTCTTATAGGGGAGA 2142
Db 2102 TACTCTCTTGGATCTTTAAAGATGGAGATACAGATATTCAGATCTCTTATAGGGGAGA 2161
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTACGAGAGTCAAGAAGTCTGCTTC 2202
Db 2162 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTACGAGAGTCAAGAAGTCTGCTTC 2221
QY 2203 TCTGTAATATGTAATTCGCGGATACCAAGCAGACATTCACACCTTTACATTTAGC 2262
Db 2222 TCTGTAATATGTAATTCGCGGATACCAAGCAGACATTCACACCTTTACATTTAGC 2281
QY 2263 AGCTGTTTATTAATTTAGAGTTGCAGAGTATTTGTTACACACGAGCTGATGTGA 2322
Db 2282 AGCTGTTTATTAATTTAGAGTTGCAGAGTATTTGTTACACACGAGCTGATGTGA 2341
QY 2323 TGCCCAAGACAAAGGAGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2342 TGCCCAAGACAAAGGAGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
QY 2383 TGTAGCAGCTCTACTAATAAGATTAATGCATGTGTCAATGCCAGGACAAATGGGCTTT 2442
Db 2402 TGTAGCAGCTCTACTAATAAGATTAATGCATGTGTCAATGCCAGGACAAATGGGCTTT 2461
QY 2443 CACACCTTTGCAGAGCAGCCCAAGGAGCAGACAGCTTTCTGCTTGTCTGCTAGC 2502
Db 2462 CACACCTTTGCAGAGCAGCCCAAGGAGCAGACAGCTTTCTGCTTGTCTGCTAGC 2521
QY 2503 CCATGAGCTGACCCGACTCTTTAAAAATCAGGAAGCAGCAACACCTTTAGATTTAGTTTC 2562
Db 2522 CCATGAGCTGACCCGACTCTTTAAAAATCAGGAAGCAGCAACACCTTTAGATTTAGTTTC 2581
QY 2563 AGCGGATGATGTCAGGGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTGCCCTCTTG 2622
Db 2582 AGCGGATGATGTCAGGGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTGCCCTCTTG 2641
QY 2623 TTACAGCCTCAAGTCTCAATGTTGTAGAGCCCGAGGCCATTCAGATGCTCTCTC 2682

Db 2642 TTACAAGCCTCAAGTCTCAATGTTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTC 2701
QY 2683 TTCAGTCCATCTACGCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTTATCTGG 2742
Db 2702 TTCAGTCCATCTACGCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTTATCTGG 2761
QY 2743 GAGTCTTTTTCAGAACTGTCTTTCAGTAGTTAGTTCAAGTGGAAACAGAGGCTGCTCCAGTTT 2802
Db 2762 GAGTCTTTTTCAGAACTGTCTTTCAGTAGTTAGTTCAAGTGGAAACAGAGGCTGCTCCAGTTT 2821
QY 2803 GGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGAAATCTTTGG 2862
Db 2822 GGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGAAATCTTTGG 2881
QY 2863 ACTTCAGCACCTAATGATATATTTGAGAGAGACAGATCACTTTGGATGTTAGTTGA 2922
Db 2882 ACTTCAGCACCTAATGATATATTTGAGAGAGACAGATCACTTTGGATGTTAGTTGA 2941
QY 2923 GATGGGACAAAGAGCTGAAGGAGATTGGAATCAATCTTTATGGACATAGGCACAACT 2982
Db 2942 GATGGGACAAAGAGCTGAAGGAGATTGGAATCAATCTTTATGGACATAGGCACAACT 3001
QY 2983 AATTAAGAGTTCGAGAGACTTATCTCCGGACACAAGGCTTTAACCCATATTTAACTTT 3042
Db 3002 AATTAAGAGTTCGAGAGACTTATCTCCGGACACAAGGCTTTAACCCATATTTAACTTT 3061
QY 3043 GAACACCTCTGCTAGTGAACAATTTATAGATCTGCTCTGATGATAAAGAGTTTCA 3102
Db 3062 GAACACCTCTGCTAGTGAACAATTTATAGATCTGCTCTGATGATAAAGAGTTTCA 3121
QY 3103 GTCTGTGGAGGAAGAGATGCAAGGTACAGTTCGAGAGCAGAGATGGAGTCTATGAGG 3162
Db 3122 GTCTGTGGAGGAAGAGATGCAAGGTACAGTTCGAGAGCAGAGATGGAGTCTATGAGG 3181
QY 3163 TGGAACTTTCACAGATACAATATTTCTCAAGATTTCAGAAAGTTCGTAAGAAATCTG 3222
Db 3182 TGGAACTTTCACAGATACAATATTTCTCAAGATTTCAGAAAGTTCGTAAGAAATCTG 3241
QY 3223 GGAAGATACACTCACCGGAGAAAAGTTCGTAAGAAATCTGTAAGAAATCTGTAAGAA 3282
Db 3242 GGAAGATACACTCACCGGAGAAAAGTTCGTAAGAAATCTGTAAGAAATCTGTAAGAA 3301
QY 3283 ACGAATGCTATTTTCATGSGTCTCCTTTTGTGAATGCAATATTCACAAAGCTTTTGATGA 3342
Db 3302 ACGAATGCTATTTTCATGSGTCTCCTTTTGTGAATGCAATATTCACAAAGCTTTTGATGA 3361
QY 3343 AAGCATCGCTACATAGTGTATGTTTGGAGCTGGCATTTTATTTGCTGAAAACCTCTTC 3402
Db 3362 AAGCATCGCTACATAGTGTATGTTTGGAGCTGGCATTTTATTTGCTGAAAACCTCTTC 3421
QY 3403 CAAAAGCAATCAATATGATATGGAATTTGGAGGAGTACTGGGTGTCAGTTCACAAGA 3462
Db 3422 CAAAAGCAATCAATATGATATGGAATTTGGAGGAGTACTGGGTGTCAGTTCACAAGA 3481
QY 3463 CAGATCTTGTACATTTGCCACAGGAGCTCTCTTTTGGCGGGTAACTTTGGGAAAAGTC 3522
Db 3482 CAGATCTTGTATACATTTGCCACAGGAGCTCTCTTTTGGCGGGTAACTTTGGGAAAAGTC 3541
QY 3523 TTTCTCTGAGTTCAGTGCATGGAATGGACATTCCTCCAGGTCATCACTCAGTCAAC 3582
Db 3542 TTTCTCTGAGTTCAGTGCATGGAATGGACATTCCTCCAGGTCATCACTCAGTCAAC 3601
QY 3583 TGGTAGGCCCCAGTCTAAATGGCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGACA 3642
Db 3602 TGGTAGGCCCCAGTCTAAATGGCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGACA 3661
QY 3643 GGCTTATCTCTGAGTATTTAAATTTACTTACCAGATTTAGGCGCTGAAGGTATGCTCGATGG 3702
Db 3662 GGCTTATCTCTGAGTATTTAAATTTACTTACCAGATTTAGGCGCTGAAGGTATGCTCGATGG 3721
QY 3703 ATAAATAGTATTTTAAAGAACTAAATTCCTCACTGAACCTAAATCATCAAGCAGCAGTGG 3762
Db 3722 ATAAATAGTATTTTAAAGAACTAAATTCCTCACTGAACCTAAATCATCAAGCAGCAGTGG 3781

QY 3763 CCTCTACGTTTACTCCTTTGCTGAAAAA 3797
 DB 3782 CCTCTACGTTTACTCCTTTGCTGAAAAA 3816

RESULT 3
 AAC66825
 ID AAC66825 standard; cDNA; 4275 BP.
 XX AAC66825;
 AC
 XX
 DT 27-FEB-2001 (first entry)
 DE Human tankyrase II coding sequence SEQ ID NO: 5.
 XX Human; tankyrase II; telomere length; signal transduction; ss.
 KW Homo sapiens.
 XX
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 284..3784
 FT /*tag= a
 FT /product= "tankyrase II"
 FT /transl_except= (pos:655..660,aa:TyrThr)
 XX
 PN WO200061813-A1.
 XX
 PD 19-OCT-2000.
 XX
 PF 10-APR-2000; 2000WO-US09558.
 XX
 PR 09-APR-1999; 99US-0128577.
 PR 13-APR-1999; 99US-0129123.
 XX
 XX (GERO-) GERON CORP.
 PA
 XX
 XX Morin GB, Funk WD, Piatyszek MA;
 PI
 XX
 DR WPI; 2000-679503/66.
 DR P-PSDB; AAB27211.
 XX
 PT Novel mammalian Tankyrase II polypeptide and the polynucleotide
 PT encoding the polypeptide useful for modulating or maintaining telomere
 PT length, replicative capacity, apoptosis, chromosome packing or gene
 PT expression
 XX
 PS Claim 1; Fig 4; 52pp; English.
 XX
 CC The present sequence is a version of the human tankyrase II coding
 CC sequence. Its protein is thought to be involved in signal transduction in
 CC the cell, and to have binding activity for other telomere-associated
 CC proteins. It is possible that it plays a role in the regulation of
 CC telomere length, thus affecting the replicative ability of the cell. The
 CC protein is useful for ribosylating target proteins, for determining
 CC tankyrase II binding activity in a sample, and for modulating telomere
 CC length in a cell.
 XX
 SQ Sequence 4275 BP; 1247 A; 889 C; 1025 G; 1114 T; 0 other;

Query Match 89.2%; Score 3388.8; DB 21; Length 4275;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 403 AGCTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAAGTGTCCCAAGC 462
 DB 481 AGCTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAAGTGTCCCAAGC 540
 QY 463 ACCTGATGATGGGGCTTATTCCTCTTCATATGATGCTCTTTGGTGCATGCTGAAGT 522
 DB 541 ACCTGATGATGGGGCTTATTCCTCTTCATATGATGCTCTTTGGTGCATGCTGAAGT 600
 QY 523 AGTCAATCTCTTTTCGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATAC 582

DB 601 AGTCAATCTCTTTTGGCGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATAC 660
 QY 583 TCCTCTCCATGAAGCTGCAATTAAGAGAAAGATTCATGTTTGCATTGTCTGTACAGCA 642
 DB 661 TCCTCTCCATGAAGCTGCAATTAAGAGAAAGATTCATGTTTGCATTGTCTGTACAGCA 720
 QY 643 TGGAGCTGAGCCAAACCATCCGAAATACAGATGGAAGCAGCAGCATTTGGATTAGCAGATCC 702
 DB 721 TGGAGCTGAGCCAAACCATCCGAAATACAGATGGAAGCAGCAGCATTTGGATTAGCAGATCC 780
 QY 703 ATCTGCCAAGCAGTGTCTTACTTGGTCAATATATAAGAAAGATGAACCTCTTGAAGTCCAG 762
 DB 781 ATCTGCCAAGCAGTGTCTTACTGGTGAATATAAGAAAGATGAACCTCTTGAAGTCCAG 840
 QY 763 GAGTGGCAATGAAGAAAAAATGATGGCTTACTACACCATTAATAATGTCAACTGCCACGC 822
 DB 841 GAGTGGCAATGAAGAAAAAATGATGGCTTACTACACCATTAATAATGTCAACTGCCACGC 900
 QY 823 AAGTATGCGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATAAAGATTAAGAT 882
 DB 901 AAGTATGCGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATAAAGATTAAGAT 960
 QY 883 TGTACAGCTGTCTTACTGCAACATGGAGCTGATGTCATGCTTAAAGATAAAGGTGATCTGCT 942
 DB 961 TGTACAGCTGTCTTACTGCAACATGGAGCTGATGTCATGCTTAAAGATAAAGGTGATCTGCT 1020
 QY 943 ACCATTACACAATGCTGTCTTATGTCATTATGAAGTAACATGAACCTTTTGGTCAAGCA 1002
 DB 1021 ACCATTACACAATGCTGTCTTATGTCATTATGAAGTAACATGAACCTTTTGGTCAAGCA 1080
 QY 1003 TGGTGCCTGTGTAATGCAATGGAGCTTGGCAATTCACCTCTTTCATGAGCAGCCTTC 1062
 DB 1081 TGGTGCCTGTGTAATGCAATGGAGCTTGGCAATTCACCTCTTTCATGAGCAGCCTTC 1140
 QY 1063 TAAGAACAGGTTGAAGTATGTTCTCTCTTAAAGTTATGTGTGACAGCCCAACTGCT 1122
 DB 1141 TAAGAACAGGTTGAAGTATGTTCTCTCTTAAAGTTATGTGTGACAGCCCAACTGCT 1200
 QY 1123 CAATTCTCAATTAAGAGTGTATAGACTTGGCTCCACACACAGTTAAAGAAAGATT 1182
 DB 1201 CAATTCTCAATTAAGAGTGTATAGACTTGGCTCCACACACAGTTAAAGAAAGATT 1260
 QY 1183 AGCATATGAATTAAGAGCCACTGTTGCTGCAAGCTGCACGAGAGCTGATGTACTCG 1242
 DB 1261 AGCATATGAATTAAGAGCCACTGTTGCTGCAAGCTGCACGAGAGCTGATGTACTCG 1320
 QY 1243 AATCAAAAAACATCTCTCTGGAATTTCAAGCATCTCTCAAAACACATGAAC 1302
 DB 1321 AATCAAAAAACATCTCTCTGGAATTTCAAGCATCTCTCAAAACACATGAAC 1380
 QY 1303 AGCATTTGCATTGTGCTGCATCTCCATATCCCAAGAAAGCAATATGTGAATGTT 1362
 DB 1381 AGCATTTGCATTGTGCTGCATCTCCATATCCCAAGAAAGCAATATGTGAATGTT 1440
 QY 1363 GCTAAAGAAAGAGCAACACATCAATGAAAGACTTAAGAAATCTTGAATCTCTGACCT 1422
 DB 1441 GCTAAAGAAAGAGCAACACATCAATGAAAGACTTAAGAAATCTTGAATCTCTGACCT 1500
 QY 1423 GGCATCTGAGAAAGCTCATATGATGTTTGAAGTAGTGGTGAACATGAAGCAAGGT 1482
 DB 1501 GGCATCTGAGAAAGCTCATATGATGTTTGAAGTAGTGGTGAACATGAAGCAAGGT 1560
 QY 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTACACAGAGCTGCATATTTGGTCACT 1542
 DB 1561 TAATGCTCTGGATAATCTTGGTCAGACTTCTACACAGAGCTGCATATTTGGTCACT 1620
 QY 1543 ACAAACTGCGGCTACTCTGAGCTATGGTGTGATCTTAACATTAATATCCCTTCAGGG 1602
 DB 1621 ACAAACTGCGGCTACTCTGAGCTATGGTGTGATCTTAACATTAATATCCCTTCAGGG 1680
 QY 1603 CTTTACTGCTTTACAGATGGAAATGAAAATGTACAGCAACTCTCTCAAGAGGATATCTC 1662

RESULT 4
AAC85294
ID AAC85294 standard; cDNA: 6019 BP.
XX
AC AAC85294;
XX
DT 29-MAR-2001 (first entry)
XX
DE Human SPANK cDNA.
XX
KW SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW poly adenosine diphosphate-ribose polymerase; catalytic domain;
KW ANK; ankyrin repeat; cytosol; insulin-responsive aminopeptidase;
KW IRAP; GLUT4; adipocyte; insulin signalling pathway; hyperlipidaemia;
KW glucose intolerance; atherosclerotic disease; atherosclerosis;
KW obesity; cardiac insufficiency; coronary insufficiency; stroke;
KW high blood pressure; non-insulin dependent diabetes; hypertension;
KW hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CD 252..3735
FT /*tag= a
FT /product= "Human SPANK"
FT /transl_except= (pos:768..878,aa:173..209)
FT (a) a SAM (sterile alpha motif) domain;
FT /transl_except= (pos:1230..1337,aa:327..362)
FT (b) a PARP (poly adenosine diphosphate-ribose polymerase) catalytic
FT /transl_except= (pos:1374..1445,aa:375..398)
FT domain; and
FT /transl_except= (pos:1695..1823,aa:482..524)
FT (c) an ANK domain composed of ankyrin repeats.
XX
PN WO200077225-A1.
XX
XX
PD 21-DEC-2000.
XX
PF 09-JUN-2000; 2000WO-US15926.
XX
PR 11-JUN-1999; 99US-0138957.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Chi N, Lodish HF;
XX
DR WPI: 2001-091404/10.
DR P-PSDB; AAB47022.
XX
PT New insulin signalling protein SPANK, useful for reducing body mass,
PT glucose intolerance or insulin resistance and for preventing or
PT treating obesity-related and muscle-related diseases -
XX
PS Claim 3; Fig 4; 65pp; English.
XX
XX This sequence represents the human SPANK cDNA. The SPANK protein
CC comprises 3 domains:
CC (a) a SAM (sterile alpha motif) domain;
CC (b) a PARP (poly adenosine diphosphate-ribose polymerase) catalytic
CC domain; and
CC (c) an ANK domain composed of ankyrin repeats.
CC SPANK is a cytosolic protein which can poly(ADP-ribose)ylate itself.
CC SPANK binds insulin-responsive aminopeptidase (IRAP) and modulates
CC translocation of GLUT4 in the perinuclear region of adipocytes. It
CC is an effector in the insulin signalling pathway in eukaryotic cells.
CC SPANK is useful for reducing body mass, reducing glucose
CC intolerance or insulin resistance, for preventing or treating
CC obesity-related diseases or disorders, such as obesity, cardiac
CC insufficiency, coronary insufficiency, stroke, hypertension,
CC atherosclerotic disease, atherosclerosis, high blood pressure, non-insulin
CC dependent diabetes, hyperlipidaemia, hyperuricemia and Syndrome X and is
CC also useful for preventing or treating muscle-related diseases or
CC disorders, such as muscular dystrophy, muscle atrophy and muscle
CC fatigue. Antibodies immunospecific for SPANK are useful for detecting

CC the presence of SPANK polypeptide in a biological sample.
XX
SQ Sequence 6019 BP; 1732 A; 1173 C; 1341 G; 1773 T; 0 other;
Query Match 89.2%; Score 3388.8; DB 22; Length 6019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 403 AGGTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAATGCTGCAAGTGTCCAAGC 462
DB |||||||
QY 449 AGGTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAATGCTGCAAGTGTCCAAGC 508
DB |||||||
QY 463 ACGTGATGATGGGCGCTTATTCCTCTTCATAATGCATGCTCTTTTGGTTCATGCTGAAGT 522
DB |||||||
QY 509 ACGTGATGATGGGCGCTTATTCCTCTTCATAATGCATGCTCTTTTGGTTCATGCTGAAGT 568
DB |||||||
QY 523 AGTCAATCTCTTTTGGCAGCATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
DB |||||||
QY 569 AGTCAATCTCTTTTGGCAGCATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 628
DB |||||||
QY 583 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTTCATGTTTGCATTTGCTCTTTACAGCA 642
DB |||||||
QY 629 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTTCATGTTTGCATTTGCTCTTTACAGCA 688
DB |||||||
QY 643 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC 702
DB |||||||
QY 689 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC 748
DB |||||||
QY 703 ATCTGCCAAAGCAGTGTCTACTGGTGAATATAAGAAAGATGAACCTTTAGAAAGTGCACG 762
DB |||||||
QY 749 ATCTGCCAAAGCAGTGTCTACTGGTGAATATAAGAAAGATGAACCTTTAGAAAGTGCACG 808
DB |||||||
QY 763 GAGTGCCCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAATGTCACCTGCCACGC 822
DB |||||||
QY 809 GAGTGCCCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAATGTCACCTGCCACGC 868
DB |||||||
QY 823 AAGTGATGCGAGAAAGTCAACTCCATTACATTTGCGACAGGATATAACAGAGTAAAGAT 882
DB |||||||
QY 869 AAGTGATGCGAGAAAGTCAACTCCATTACATTTGCGACAGGATATAACAGAGTAAAGAT 928
DB |||||||
QY 883 TGTACAGCTGTACTGCAACATGGAGCTGATGCTTAAAGATAAAGTGTGCTGTGT 942
DB |||||||
QY 929 TGTACAGCTGTACTGCAACATGGAGCTGATGCTTAAAGATAAAGTGTGCTGTGT 988
DB |||||||
QY 943 ACCATTACACAATGCTGTCTTATGGTCATTATCAAGTAACTGAACCTTTTGGTCAAGCA 1002
DB |||||||
QY 989 ACCATTACACAATGCTGTCTTATGGTCATTATCAAGTAACTGAACCTTTTGGTCAAGCA 1048
DB |||||||
QY 1003 TGGTGCCCTGTGAATGCAATGGAGCTTGGGCAATTCACCTCTCTTCATGAGGCGAGCTTC 1062
DB |||||||
QY 1049 TGGTGCCCTGTGAATGCAATGGAGCTTGGGCAATTCACCTCTCTTCATGAGGCGAGCTTC 1108
DB |||||||
QY 1063 TAAGAACAGGGTTGAAGTATGTTCTTCTTAAAGTTATGGTGCAGAGCCCAACACTGCT 1122
DB |||||||
QY 1109 TAAGAACAGGGTTGAAGTATGTTCTTCTTAAAGTTATGGTGCAGAGCCCAACACTGCT 1168
DB |||||||
QY 1123 CAATTGTCCACAATAAAGTGTCTATAGACTTGGCTCCCAACACACAGTTAAAAAGAAAGATT 1182
DB |||||||
QY 1169 CAATTGTCCACAATAAAGTGTCTATAGACTTGGCTCCCAACACACAGTTAAAAAGAAAGATT 1228
DB |||||||
QY 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAGCTGTGTTACTCG 1242
DB |||||||
QY 1229 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAGCTGTGTTACTCG 1288
DB |||||||
QY 1243 AATCAAAAACATCTCTCTCGGAAATGGTGAATTTCAAGCATCTCAACACATGAAC 1302
DB |||||||
QY 1289 AATCAAAAACATCTCTCTCGGAAATGGTGAATTTCAAGCATCTCAACACATGAAC 1348
DB |||||||
QY 1303 AGCATTGCATTTGCTGCTGCTATCTCCATATCCCAAAAAGAAAGAAATATGGAAGTGT 1362
DB |||||||
QY 1349 AGCATTGCATTTGCTGCTGCTATCTCCATATCCCAAAAAGAAAGAAATATGGAAGTGT 1408
DB |||||||
QY 1363 GCTAAGAAAAGAGCAACATCAATGAAAGACTTAAGAAATTTCTTGACTCTCTGCACGT 1422

|||||
Db 1409 GCTAGAAAAAGCAACATCAATGAAAACATTAAGAATCTTGACTCTCTGACAGT 1468
Qy 1423 GGCACTCAGAAAAGCTCAATAATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGGT 1482
Db 1469 GGCATCTCAGAAAAGCTCAATAATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGGT 1528
Qy 1483 TAATGCTCTGATATCTTTGGTCAGACTTCTCTACAGAGCTGCATATTTGTCATCT 1542
Db 1529 TAATGCTCTGATATCTTTGGTCAGACTTCTCTACAGAGCTGCATATTTGTCATCT 1588
Qy 1543 ACAACCTGCGCGCTACTCTGAGCTATGGGTGATGCTTCAACATATATCCCTTCAGGG 1602
Db 1589 ACAACCTGCGCGCTACTCTGAGCTATGGGTGATGCTTCAACATATATCCCTTCAGGG 1648
Qy 1603 CTTTACTGCTTTACAGATGGGAATGAAAATGTACAGCAACTCTCCAGAGGTATCTC 1662
Db 1649 CTTTACTGCTTTACAGATGGGAATGAAAATGTACAGCAACTCTCCAGAGGTATCTC 1708
Qy 1663 ATTAGGTAATTCAGAGCAGACAGACAATGCTGGAAGCTCAAGGCTGGAGATGTGGA 1722
Db 1709 ATTAGGTAATTCAGAGCAGACAGACAATGCTGGAAGCTCAAGGCTGGAGATGTGGA 1768
Qy 1723 AACTGTAAAAAATCTGTACTGTTTCAGAGTGTCAACTSCAGAGACATTTGAAGGGCGCTCA 1782
Db 1769 AACTGTAAAAAATCTGTACTGTTTCAGAGTGTCAACTSCAGAGACATTTGAAGGGCGCTCA 1828
Qy 1783 GTCTACACACTTCAATTTTGACGTGGGTATTAACAGAGTGTCCGTGGGAATATCTGCT 1842
Db 1829 GTCTACACACTTCAATTTTGACGTGGGTATTAACAGAGTGTCCGTGGGAATATCTGCT 1888
Qy 1843 ACAGCATGAGCTGATGTGCTGCTAAAGATAAAGAGGCGCTGTGACCTTTGCACAAATGC 1902
Db 1889 ACAGCATGAGCTGATGTGCTGCTAAAGATAAAGAGGCGCTGTGACCTTTGCACAAATGC 1948
Qy 1903 ATGTTCTTATGAGACATTAATGAAGTTGCCAGAACTTCTGTTAAACATGAGGAGTAGTTAA 1962
Db 1949 ATGTTCTTATGAGACATTAATGAAGTTGCCAGAACTTCTGTTAAACATGAGGAGTAGTTAA 2008
Qy 1963 TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGACAGCAAAAGGAATATGA 2022
Db 2009 TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGACAGCAAAAGGAATATGA 2068
Qy 2023 AATTTGCAAACTCTGCTCCAGCATGGTGCAGACCTTACCAGCAAAACAGGATGGAAA 2082
Db 2069 AATTTGCAAACTCTGCTCCAGCATGGTGCAGACCTTACCAGCAAAACAGGATGGAAA 2128
Qy 2083 TACTCCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGGAGA 2142
Db 2129 TACTCCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGGAGA 2188
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAGTCTGCTTC 2202
Db 2189 TGCAAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAGTCTGCTTC 2248
Qy 2203 TCTGTATAATGTAATTTGCCGGATACCCAGGCGAGACATTCACACCTTTACATTTAGC 2262
Db 2249 TCTGTATAATGTAATTTGCCGGATACCCAGGCGAGACATTCACACCTTTACATTTAGC 2308
Qy 2263 AGCTGGTTATATAATTTAGAGTTGACAGATATTTGTTTACAAACAGGAGCTGATGTGAA 2322
Db 2309 AGCTGGTTATATAATTTAGAGTTGACAGATATTTGTTTACAAACAGGAGCTGATGTGAA 2368
Qy 2323 TGCCCCAAGCAAGAGGAGCTTATCTCTTACATATGACAGATCTTACGGGCGATGTAGA 2382
Db 2369 TGCCCCAAGCAAGAGGAGCTTATCTCTTACATATGACAGATCTTACGGGCGATGTAGA 2428
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATTAATGCATGTGTCAATGCCAGGCAAAATGGGCTTT 2442
Db 2429 TGTAGCAGCTCTACTAATAAAGTATTAATGCATGTGTCAATGCCAGGCAAAATGGGCTTT 2488
Qy 2443 CACACCTTTGCGAAGACAGCCCAAAAGGAGGACACAGCTTTGTGCTTTGCTGCTAC 2502
|||||

Db 2489 CACACCTTTGCGAAGACAGCCCAAAAGGAGCAGACACAGCTTTGTGCTTTGTGCTAGC 2548
Qy 2503 CCATGAGCTGACCCGACTCTTAAATCAGGAAGCAGACAAACACCTTTAGATTTAGTTTC 2562
Db 2549 CCATGAGCTGACCCGACTCTTAAATCAGGAAGCAGACAAACACCTTTAGATTTAGTTTC 2608
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCATGCCCCATCTGCTGCCCCCTTG 2622
Db 2609 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCATGCCCCATCTGCTGCCCCCTTG 2668
Qy 2623 TTACAAGCCTCAAGTGTCTCAATGGTGTGAGAAGCCAGAGCCACTGCAAGTCTCTC 2682
Db 2669 TTACAAGCCTCAAGTGTCTCAATGGTGTGAGAAGCCAGAGCCACTGCAAGTCTCTC 2728
Qy 2683 TTCAAGTTCATCTAGCCCATCAAGCCTTTCTGAGCAGCAGCTTTTGACAACCTATCTGG 2742
Db 2729 TTCAAGTTCATCTAGCCCATCAAGCCTTTCTGAGCAGCAGCTTTTGACAACCTATCTGG 2788
Qy 2743 GAGTTTTTCAGAACTGTCTTCAAGTGTAGTTAGTTCAAGTGGAAACAGAGGGTCTTCCAGTTT 2802
Db 2789 GAGTTTTTCAGAACTGTCTTCAAGTGTAGTTAGTTCAAGTGGAAACAGAGGGTCTTCCAGTTT 2848
Qy 2803 GGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAGGAATCTTGG 2862
Db 2849 GGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAGGAATCTTGG 2908
Qy 2863 ACTTGACACCTAATGATATATTTGAGAGAGAACAGATCACTTTGGATGATATTAGTTGA 2922
Db 2909 ACTTGACACCTAATGATATATTTGAGAGAGAACAGATCACTTTGGATGATATTAGTTGA 2968
Qy 2923 GATGGGACACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACT 2982
Db 2969 GATGGGACACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACT 3028
Qy 2983 AATTAAGGAGTGCAGAGACTTATCTCCGACACACAGGCTTTAACCCATATTTAACTTT 3042
Db 3029 AATTAAGGAGTGCAGAGACTTATCTCCGACACACAGGCTTTAACCCATATTTAACTTT 3088
Qy 3043 GAACACCTCTGGTAGTGGAAACAATTTATAGATCTCTCTCCTGATGATAAGAGTTTCA 3102
Db 3089 GAACACCTCTGGTAGTGGAAACAATTTATAGATCTCTCTCCTGATGATAAGAGTTTCA 3148
Qy 3103 GTCTGTGAGAGAGATGCAAGTACAGTTCCGAGACACAGAGATGGAGTCTATGCGAG 3162
Db 3149 GTCTGTGAGAGAGATGCAAGTACAGTTCCGAGACACAGAGATGGAGTCTATGCGAG 3208
Qy 3163 TGGAACTTCAACAGATACAAATTTCTCAAGATTCAGAAGTTTGTAAACAAGAACTATG 3222
Db 3209 TGGAACTTCAACAGATACAAATTTCTCAAGATTCAGAAGTTTGTAAACAAGAACTATG 3268
Qy 3223 GGAAGATACACTCACCAGGAGAAAGTTCCTGAAGAAAACCAACCACTGCAATGA 3282
Db 3269 GGAAGATACACTCACCAGGAGAAAGTTCCTGAAGAAAACCAACCACTGCAATGA 3328
Qy 3283 ACGAATGCTATTTGATGGGCTCTCTTTTGGAATGCAATTTCCACAAAGGCTTTGATGA 3342
Db 3329 ACGAATGCTATTTGATGGGCTCTCTTTTGGAATGCAATTTCCACAAAGGCTTTGATGA 3388
Qy 3343 AAGCATGCTGATGATGATGTTTGGAGTGGCAATTTATTTGCTGAAAACCTCTTC 3402
Db 3389 AAGCATGCTGATGATGATGTTTGGAGTGGCAATTTATTTGCTGAAAACCTCTTC 3448
Qy 3403 CAAAAGCAATCAATATGATATGGAATTTGGAGGAGTCTGGGTGTCAGTTCAACAAGA 3462
Db 3449 CAAAAGCAATCAATATGATATGGAATTTGGAGGAGTCTGGGTGTCAGTTCAACAAGA 3508
Qy 3463 CAGATCTTTGATTTGCCACAGGAGCTCTCTTTTGGCGGGTAACCTTTGGGAAAGTC 3522
Db 3509 CAGATCTTTGATTTGCCACAGGAGCTCTCTTTTGGCGGGTAACCTTTGGGAAAGTC 3568
Qy 3523 TTTCTGCGAGTTGCAATGAAAATGGCAATTTCTCTCCTCAGGTCATCACTCAGTCAAC 3582
Db 3569 TTTCTGCGAGTTGCAATGAAAATGGCAATTTCTCTCCTCAGGTCATCACTCAGTCAAC 3628

Db	1348	ACAAACCTGCGCGCTACTCTCTGAGCTATGGGTGTGATCTCTAAACATTATATCCCTTCAGGG	1407
QY	1603	CTTTACTGCTTTACAGATGGGAATGAAATATGTACAGCAACTCTCCCAAGAGGATCTC	1662
Db	1408	CTTTACTGCTTTACAGATGGGAATGAAATATGTACAGCAACTCTCCCAAGAGGATCTC	1467
QY	1663	ATTAGGTAAATCAGAGCGACAGACAATGTCTGGAAGTGCAAAGCTGGAGATGTGCA	1722
Db	1468	ATTAGGTAAATCAGAGCGACAGACAATGTCTGGAAGTGCAAAGCTGGAGATGTGCA	1527
QY	1723	AACTGTAAAAAACTGTACTCTTCAGAGTGTCAACTGCCAGAGACATTGAAGGCGGTCA	1782
Db	1528	AACTGTAAAAAACTGTACTCTTCAGAGTGTCAACTGCCAGAGACATTGAAGGCGGTCA	1587
QY	1783	GTCTACACCACTTCATTTTGCAGCTGGGTATTAACAGAGTGTCCGTGGTGAATATCTGCT	1842
Db	1588	GTCTACACCACTTCATTTTGCAGCTGGGTATTAACAGAGTGTCCGTGGTGAATATCTGCT	1647
QY	1843	ACAGCATGGAGCTGATGTGCATGCTTAACAGTAAAGAGAGGCGCTTGTACCTTTGCACAAATGC	1902
Db	1648	ACAGCATGGAGCTGATGTGCATGCTTAACAGTAAAGAGAGGCGCTTGTACCTTTGCACAAATGC	1707
QY	1903	ATGTTCTTATGGACATTATGAAGCTTCGACAACTCTTCTTAAACATGGAGCAGTAGTTAA	1962
Db	1708	ATGTTCTTATGGACATTATGAAGCTTCGACAACTCTTCTTAAACATGGAGCAGTAGTTAA	1767
QY	1963	TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAAAAGGAAATATGA	2022
Db	1768	TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAAAAGGAAATATGA	1827
QY	2023	AATTTGCAAACTTCTGCTCCAGCATGGTGACAGCCCTACCAAAAACAGGGATGGAAA	2082
Db	1828	AATTTGCAAACTTCTGCTCCAGCATGGTGACAGCCCTACCAAAAACAGGGATGGAAA	1887
QY	2083	TACTCCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTCAGATCTGCTTTAGGGGAGA	2142
Db	1888	TACTCCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTCAGATCTGCTTTAGGGGAGA	1947
QY	2143	TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTACGAGAGTGGAAGTTGCTTCTTC	2202
Db	1948	TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTACGAGAGTGGAAGTTGCTTCTTC	2007
QY	2203	TCCTGTAAATGTAATTTGCCGGGATACCCAGGGCAGACATTCACACCTTTACATTTAGC	2262
Db	2008	TCCTGTAAATGTAATTTGCCGGGATACCCAGGGCAGACATTCACACCTTTACATTTAGC	2067
QY	2263	AGCTGGTTATAATTTAGAAAGTTGCAGAGTATTTGTTACACACGGAGCTCATGTGAA	2322
Db	2068	AGCTGGTTATAATTTAGAAAGTTGCAGAGTATTTGTTACACACGGAGCTCATGTGAA	2127
QY	2323	TGCCCCAAGACAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCGATGTAGA	2382
Db	2128	TGCCCCAAGACAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCGATGTAGA	2187
QY	2383	TGTAGCAGCTCTACTAATAAGTATAATGCATGTGTCAATGCCACGGACAAATGGCGTTT	2442
Db	2188	TGTAGCAGCTCTACTAATAAGTATAATGCATGTGTCAATGCCACGGACAAATGGCGTTT	2247
QY	2443	CACACCTTTGCCAGCAGCGCCAAAAGGGAGCAACACAGCTTTGTCTTTGTGTGCTAGC	2502
Db	2248	CACACCTTTGCCAGCAGCGCCAAAAGGGAGCAACACAGCTTTGTGTGCTAGC	2307
QY	2503	CCATGAGCTGACCGACTCTTAAATTCAGGAAGGACAAACACCTTTAGATTTAGTTTTC	2562
Db	2308	CCATGAGCTGACCGACTCTTAAATTCAGGAAGGACAAACACCTTTAGATTTAGTTTTC	2367
QY	2563	AGCGGATGATGTACGCGTCTTCTGACAGAGCCATGCCGCCATCTGCTCTGCCCTCTTTG	2622
Db	2368	AGCAGATGATGTACGCGTCTTCTGACAGAGCCATGCCGCCATCTGCTCTGCCCTCTTTG	2427
QY	2623	TTACAAGCCTCAAGTGTCTCAATGTGTGAGAAGCCAGGACCACTGCAGATGCTCTCTC	2682

QY 3763 CCTCTACGTTTACTCCTTTGCTGAAAAAAA 3794
 |||||||||||||||||||||||||||||||||||
 Db 3568 CCTCTACGTTTACTCCTTTGCTGAAAAAAA 3599
 |||||||||||||||||||||||||||||||||||

RESULT 6
 AAF63930
 ID AAF63930 standard; DNA; 4406 BP.
 AC AAF63930;
 XX
 DT
 XX 05-APR-2001 (first entry)
 XX Human tankyrase2 clone consensus oligonucleotide SEQ ID NO: 106.
 DE
 XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
 KW inflammatory disorder; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200100849-A1.
 XX
 XX 04-JAN-2001.
 XX 28-JUN-2000; 2000WO-US17827.
 XX 29-JUN-1999; 99US-0141582.
 XX {ICOS-} ICOS CORP.
 XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
 PI
 XX WPI; 2001-102896/11.
 DR P-PSDB; AAB66290.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
 PT inflammatory and autoimmune disorders -
 XX
 PS Example 2; Page 168-173; 242pp; English.

XX The present invention provides the protein and coding sequence for the
 CC human tankyrase2 protein. This is found in two different versions,
 CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
 CC polyADP-ribosylation activity and is involved in the modification of
 CC TRF1, which is a telomere-specific binding protein. The regulation of
 CC telomere length, in which TRF1 has a role, is linked to ageing and
 CC cancer. The sequences are useful in the treatment of cancers and
 CC inflammatory disorders.

XX
 XX Sequence 4406 BP; 1268 A; 910 C; 1054 G; 1174 T; 0 other;
 SQ

Query Match 89.2%; Score 3387.2; DB 22; Length 4406;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 403 AGGTTTGGGCGGAAAGACGCTAGTTGAATATTGCTTCAGAAATGGTGCAAGTGCCCAAGC 462
 |||||||||||||||||||||||||||||||||||
 Db 487 AGGTTTGGGCGGAAAGACGCTAGTTGAATATTGCTTCAGAAATGGTGCAAGTGCCCAAGC 546
 |||||||||||||||||||||||||||||||||||

QY 463 ACGTGATGATGGGGCCCTTATCTCTTCATPAATGCATGCTTTTGGTCATGCTGAAGT 522
 |||||||||||||||||||||||||||||||||||
 Db 547 ACGTGATGATGGGGCCCTTATCTCTTCATPAATGCATGCTTTTGGTCATGCTGAAGT 606
 |||||||||||||||||||||||||||||||||||

QY 523 AGTCAATCTCTTTGGACATGTTGCAGACCCCAATGCTCGAGATTAATGGAATTATAC 582
 |||||||||||||||||||||||||||||||||||
 Db 607 AGTCAATCTCTTTGGACATGTTGCAGACCCCAATGCTCGAGATTAATGGAATTATAC 666
 |||||||||||||||||||||||||||||||||||

QY 583 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTGATGTTTGGCATTTGCTGTTTACAGCA 642
 |||||||||||||||||||||||||||||||||||
 Db 667 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTGATGTTTGGCATTTGCTGTTTACAGCA 726
 |||||||||||||||||||||||||||||||||||

QY 643 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGACAGCAGCAATTTAGCAGATCC 702
 |||||||||||||||||||||||||||||||||||
 Db 727 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGACAGCAGCAATTTAGCAGATCC 786
 |||||||||||||||||||||||||||||||||||

QY 703 ATCTGCCAAAGCAGTCTTACTGGTCAATATAGAAGATGAACCTTTAGAAAAGTCCAG 762
 |||||||||||||||||||||||||||||||||||
 Db 787 ATCTGCCAAAGCAGTCTTACTGGTCAATATAGAAGATGAACCTTTAGAAAAGTCCAG 846
 |||||||||||||||||||||||||||||||||||

QY 763 GAGTGGCAATGAAGAAAAAATGATGCTTACTCACACCATTTAAATGTCAACTGCCACGC 822
 |||||||||||||||||||||||||||||||||||
 Db 847 GAGTGGCAATGAAGAAAAAATGATGCTTACTCACACCATTTAAATGTCAACTGCCACGC 906
 |||||||||||||||||||||||||||||||||||

QY 823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGAT 882
 |||||||||||||||||||||||||||||||||||
 Db 907 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGAT 966
 |||||||||||||||||||||||||||||||||||

QY 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCTTAAAGATAAAGTCACTCTGGT 942
 |||||||||||||||||||||||||||||||||||
 Db 967 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCTTAAAGATAAAGTCACTCTGGT 1026
 |||||||||||||||||||||||||||||||||||

QY 943 ACCATTACACAATGCCTGTTCTTATGCTCATTTAAGTAACCTGAACCTTTTGGTCAAGCA 1002
 |||||||||||||||||||||||||||||||||||
 Db 1027 ACCATTACACAATGCCTGTTCTTATGCTCATTTAAGTAACCTGAACCTTTTGGTCAAGCA 1086
 |||||||||||||||||||||||||||||||||||

QY 1003 TGTGCTGTGTAATGCAATGCAATGCAATGCGCAATTCACCTCTTTCATGAGGAGCTTC 1062
 |||||||||||||||||||||||||||||||||||
 Db 1087 TGTGCTGTGTAATGCAATGCAATGCAATGCGCAATTCACCTCTTTCATGAGGAGCTTC 1146
 |||||||||||||||||||||||||||||||||||

QY 1063 TAAGAACAGGGTTGAAGTATGTTCTCTCTTAAAGTTATGTTGCAGACCCCAACACTGCT 1122
 |||||||||||||||||||||||||||||||||||
 Db 1147 TAAGAACAGGGTTGAAGTATGTTCTCTCTTAAAGTTATGTTGCAGACCCCAACACTGCT 1206
 |||||||||||||||||||||||||||||||||||

QY 1123 CAATTGTCACAATAAAGTGTCTATAGACTTGGCTCCACACCACAGTTAAAGAAAGATT 1182
 |||||||||||||||||||||||||||||||||||
 Db 1207 CAATTGTCACAATAAAGTGTCTATAGACTTGGCTCCACACCACAGTTAAAGAAAGATT 1266
 |||||||||||||||||||||||||||||||||||

QY 1183 AGCATGTAATTAAGGGCCACTGCTGTCGCAAGCTGCAAGCTGCAAGGAGCACTGATTACTCG 1242
 |||||||||||||||||||||||||||||||||||
 Db 1267 AGCATGTAATTAAGGGCCACTGCTGTCGCAAGCTGCAAGGAGCACTGATTACTCG 1326
 |||||||||||||||||||||||||||||||||||

QY 1243 AATCAAAAACATCTCTCTGGAATGGTCAATTTCAAGCATCTCAAAACACATGAAAC 1302
 |||||||||||||||||||||||||||||||||||
 Db 1327 AATCAAAAACATCTCTCTGGAATGGTCAATTTCAAGCATCTCAAAACACATGAAAC 1386
 |||||||||||||||||||||||||||||||||||

QY 1303 AGCATTCATTTGCTGCTGCTCATCTCCATATCCCAAAAGAAAGCAATATGTGAACCTGTT 1362
 |||||||||||||||||||||||||||||||||||
 Db 1387 AGCATTCATTTGCTGCTGCTCATCTCCATATCCCAAAAGAAAGCAATATGTGAACCTGTT 1446
 |||||||||||||||||||||||||||||||||||

QY 1363 GCTAAGAAAAGGACCAACATCAATGAAAGACTTAAAGAAATTTCTTGACTCTCTCTGCAGT 1422
 |||||||||||||||||||||||||||||||||||
 Db 1447 GCTAAGAAAAGGACCAACATCAATGAAAGACTTAAAGAAATTTCTTGACTCTCTCTGCAGT 1506
 |||||||||||||||||||||||||||||||||||

QY 1423 GGCATCTGAGAAAGCTCATATATGATTTGTTGAAGTAGTGGTGAACATGAAGCAAGGT 1482
 |||||||||||||||||||||||||||||||||||
 Db 1507 GGCATCTGAGAAAGCTCATATATGATTTGTTGAAGTAGTGGTGAACATGAAGCAAGGT 1566
 |||||||||||||||||||||||||||||||||||

QY 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACAGAGCTGCATATTTGGTGCATCT 1542
 |||||||||||||||||||||||||||||||||||
 Db 1567 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACAGAGCTGCATATTTGGTGCATCT 1626
 |||||||||||||||||||||||||||||||||||

QY 1543 ACAACCTGCGCCCTACTCTGAGCTATGGGTGATCTTAACTATATATCCCTTCAGGG 1602
 |||||||||||||||||||||||||||||||||||
 Db 1627 ACAACCTGCGCCCTACTCTGAGCTATGGGTGATCTTAACTATATATCCCTTCAGGG 1686
 |||||||||||||||||||||||||||||||||||

QY 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCTCTCCAGAGGTATCTC 1662
 |||||||||||||||||||||||||||||||||||
 Db 1687 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCTCTCCAGAGGTATCTC 1746
 |||||||||||||||||||||||||||||||||||

QY 1663 ATTAGGTAATTCAGAGGACAGACAAATTCCTGGAAGCTGCAAGGCTGGAGATGTGCA 1722
 |||||||||||||||||||||||||||||||||||
 Db 1747 ATTAGGTAATTCAGAGGACAGACAAATTCCTGGAAGCTGCAAGGCTGGAGATGTGCA 1806
 |||||||||||||||||||||||||||||||||||

QY 1723 AACTGTAAAAAAACTGTGTACTGTTCAAGTGTCAACTGCAGAGACATTTGAAGGGCGTCA 1782
 |||||||||||||||||||||||||||||||||||

[illegible]

Db	2887	GGAGAAAAGAGAGGCTCCAGGAGTATGATTTTAGCATAACTCAATTGTAAGGAATCTTGG	2946
Qy	2863	ACTTGAGCACCTAATGATATATTTTGAGAGAGAACAGATCACTTTGGATGTATAGTTGA	29222
Db	2947	ACTTGAGCACCTAATGATATATTTTGAGAGAGAACAGATCACTTTGGATGTATAGTTGA	3006
Qy	2923	GATGGGCACAAGAGCTGAAGAGATTTGGAATCAATGCTTTATGGACATAGGCACAAACT	2982
Db	3007	GATGGGCACAAGAGCTGAAGAGATTTGGAATCAATGCTTTATGGACATAGGCACAAACT	3066
Qy	2983	AATTTAAGGACTCCAGAGACTTATCTCCGGACACAAGGCTCTTAACCATATTTAACTTT	3042
Db	3067	AATTTAAGGACTCCAGAGACTTATCTCCGGACACAAGGCTCTTAACCATATTTAACTTT	3126
Qy	3043	GAACACCTCTGGTAGTGGAAACAATCTTATAGATCTGTCTCTGATGATAAAGAGTTTCA	3102
Db	3127	GAACACCTCTGGTAGTGGAAACAATCTTATAGATCTGTCTCTGATGATAAAGAGTTTCA	3186
Qy	3103	GTCTGTGGAGAGAGATGCAAAAGTACAGTTCCAGAGCACAGAGATGGAGGTCATGCAGG	3162
Db	3187	GTCTGTGGAGAGAGATGCAAAAGTACAGTTCCAGAGCACAGAGATGGAGGTCATGCAGG	3246
Qy	3163	TGGAATCTTCAACAGATACAAATATCTCAAGATTTCAGAGGTTTCTTAACAAGAAACTATG	3222
Db	3247	TGGAATCTTCAACAGATACAAATATCTCAAGATTTCAGAGGTTTCTTAACAAGAAACTATG	3306
Qy	3223	GGAAAGATACACTCACCGGAGAAAAGAGTTCTTGAAGAAAAACCAACCATGCCAATGA	3282
Db	3307	GGAAAGATACACTCACCGGAGAAAAGAGTTCTTGAAGAAAAACCAACCATGCCAATGA	3366
Qy	3283	ACGAATGCTATTTTCATGGGTCCTTTTGTGAATGCAATTTATCCAAAGGCTTTTGATGA	3342
Db	3367	ACGAATGCTATTTTCATGGGTCCTTTTGTGAATGCAATTTATCCAAAGGCTTTTGATGA	3426
Qy	3343	AAGCATCGCTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGTCTGAAAACTCTTC	3402
Db	3427	AAGCATCGCTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGTCTGAAAACTCTTC	3486
Qy	3403	CAAAAGCAATCAATATGTATATGGAATTTGGAGAGGTACTGGGTGCCAGTTCACAAAGA	3462
Db	3487	CAAAAGCAATCAATATGTATATGGAATTTGGAGAGGTACTGGGTGCCAGTTCACAAAGA	3546
Qy	3463	CAGATCTGTTTACATTTGCCAGGAGGAGCTGCTTTTCCGGGTAACTTTGGGAAAGTC	3522
Db	3547	CAGATCTGTTTACATTTGCCAGGAGGAGCTGCTTTTCCGGGTAACTTTGGGAAAGTC	3606
Qy	3523	TTTCCTCAGTTCAGTGCATGAAATGGCAATCTCTCCAGGTCACTCACTCAGTCCAC	3582
Db	3607	TTTCCTCAGTTCAGTGCATGAAATGGCAATCTCTCCAGGTCACTCACTCAGTCCAC	3666
Qy	3583	TGGTAGGCCAGTGTAAATGGCCTAGCATTTAGCTGAATATGTTATTACAGAGGAGAACA	3642
Db	3667	TGGTAGGCCAGTGTAAATGGCCTAGCATTTAGCTGAATATGTTATTATTACAGAGGAGAACA	3726
Qy	3643	GGCTTATCCTGAGTATTTTAATTAATTTACCAGATTATAGGCTCGAAGTATGGTCGATGG	3702
Db	3727	GGCTTATCCTGAGTATTTTAATTAATTTACCAGATTATAGGCTCGAAGTATGGTCGATGG	3786
Qy	3703	ATAAATAGTATTTTAAAGAACTAATTCACCTGAACTTAAATCATCAAAAGCAGCAGTGG	3762
Db	3787	ATAAATAGTATTTTAAAGAACTAATTCACCTGAACTTAAATCATCAAAAGCAGCAGTGG	3846
Qy	3763	CCCTACGTTTTTACTCCTTTTGTGAAAAAAA	3794
Db	3847	CCCTACGTTTTTACTCCTTTTGTGAAAAAAA	3878

RESULT 7
AAF63953
ID AAF63953 standard; DNA; 4992 BP.
XX
AC AAF63953;
XX

DT 05-APR-2001 (first entry)
 XX Human tankyrase2 TANK2-SHORT coding sequence SEQ ID NO: 134.
 DE Human tankyrase2 TANK2; TRF1; telomere; cancer; neoplasm; aging;
 XX Human tankyrase2 TANK2; TRF1; telomere; cancer; neoplasm; aging;
 KW inflammatory disorder; ds.
 KW inflammatory disorder; ds.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO200100849-A1.
 PN 04-JAN-2001.
 PD 28-JUN-2000; 2000WO-US17827.
 XX 29-JUN-1999; 99US-0141582.
 XX (ICOS-) ICOS CORP.
 XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
 PI WPI: 2001-102896/11.
 DR P-PSDB; AAB66295.
 XX New tankyrase2 polypeptides, useful for treating conditions mediated by
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
 PT inflammatory and autoimmune disorders.
 XX Claim 6: Page 195-200; 242pp; English.
 XX The present invention provides the protein and coding sequence for the
 CC human tankyrase2 protein. This is found in two different versions,
 CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
 CC polyADP-ribosylation activity and is involved in the modification of
 CC TRF1, which is a telomere-specific binding protein. The regulation of
 CC telomere length, in which TRF1 has a role, is linked to ageing and
 CC cancer. The sequences are useful in the treatment of cancers and
 CC inflammatory disorders.
 XX Sequence 4992 BP; 1454 A; 1064 C; 1191 G; 1283 T; 0 other;
 SQ
 Query Match 89.2%; Score 3387.2; DB 22; Length 4992;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 403 AGGTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGATGGTGGCAAGTGTCCAAGC 462
 DB 1073 AGGTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGATGGTGGCAAGTGTCCAAGC 1132
 QY 463 AGCTGATGATGGGGCTTATTCTCTTCATATATGATGCTCTTTTGGTCATGCTGAAGT 522
 DB 1133 AGCTGATGATGGGGCTTATTCTCTTCATATATGATGCTCTTTTGGTCATGCTGAAGT 1192
 QY 523 AGTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATTAATGGAATTATAC 582
 DB 1193 AGTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATTAATGGAATTATAC 1252
 QY 583 TCCTCTCCATGAGCTGCAATTAAGGAAGATTGATGTTGCAATGCTGTTACAGCA 642
 DB 1253 TCCTCTCCATGAGCTGCAATTAAGGAAGATTGATGTTGCAATGCTGTTACAGCA 1312
 QY 643 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACGAGCATTTGGATTAGCAGATCC 702
 DB 1313 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACGAGCATTTGGATTAGCAGATCC 1372
 QY 703 ATCTGCCAAGCAGTCTTACCTGGTGAATATTAAGAAAGATGAATCTTGAAGAGTCCAG 762
 DB 1373 ATCTGCCAAGCAGTCTTACCTGGTGAATATTAAGAAAGATGAATCTTGAAGAGTCCAG 1432
 QY 763 GAGTGGCAATGAAGAAATGATGCTCTACTCACACCATTAATTAATGTCACCTGCCAGC 822
 DB 1433 GAGTGGCAATGAAGAAATGATGCTCTACTCACACCATTAATTAATGTCACCTGCCAGC 1492

QY 823 AAGTGATGCGAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
 DB 1493 AAGTGATGCGAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1552
 QY 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCATGCTAAAGATAAAGAGTGAATCTGGT 942
 DB 1553 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCATGCTAAAGATAAAGAGTGAATCTGGT 1612
 QY 943 ACCATTACACAATGCTGTTCTTATGCTCATATGAAGTAAGTGAATCTTTGGTCAAGCA 1002
 DB 1613 ACCATTACACAATGCTGTTCTTATGCTCATATGAAGTAAGTGAATCTTTGGTCAAGCA 1672
 QY 1003 TGGTGCTGTGTAATGCAATGGAGCTTGTGCAATTAACCTCTTTCATGAGGACGCTTC 1062
 DB 1673 TGGTGCTGTGTAATGCAATGGAGCTTGTGCAATTAACCTCTTTCATGAGGACGCTTC 1732
 QY 1063 TAAGAACAGGGTTGAAGTATGTTCTTCTTAAGTTATGTTGCGACAGCCCAACACTGCT 1122
 DB 1733 TAAGAACAGGGTTGAAGTATGTTCTTCTTAAAGTTATGTTGCGACAGCCCAACACTGCT 1792
 QY 1123 CAATGTCACAATAAAGTGCTATAGACTTGGCTCCACACACACAGTTAAAGCAAGATTT 1182
 DB 1793 CAATGTCACAATAAAGTGCTATAGACTTGGCTCCACACACAGTTAAAGCAAGATTT 1852
 QY 1183 AGCATATGAATTTAAAGGCGCACTGCTGCAAGTGCACGAGAGAGCTGATTTACTCG 1242
 DB 1853 AGCATATGAATTTAAAGGCGCACTGCTGCAAGTGCACGAGAGAGCTGATTTACTCG 1912
 QY 1243 AATCAAAAAACATCTCTCTGGAATGGTGAATTTCAAGCATCTTCAACACATGAAC 1302
 DB 1913 AATCAAAAAACATCTCTCTGGAATGGTGAATTTCAAGCATCTTCAACACATGAAC 1972
 QY 1303 ASCATTGCAATTTGCGGCGCATCTCCATATCCCAAGAAAGCAAAATATGTCAACTGTT 1362
 DB 1973 AGCATGCAATTTGCGGCGCATCTCCATATCCCAAGAAAGCAAAATATGTCAACTGTT 2032
 QY 1363 GCTAAGAAAGAGGAGCAACATCAATGAAAGACTTAAAGAAATTTGACCTCTGACCGT 1422
 DB 2033 GCTAAGAAAGAGGAGCAACATCAATGAAAGACTTAAAGAAATTTGACCTCTGACCGT 2092
 QY 1423 GGCATCTGAGAAGCTCATATGATGTTGTAAGTAGTGGTGAACATGAAGCAAGGT 1482
 DB 2093 GGCATCTGAGAAGCTCATATGATGTTGTAAGTAGTGGTGAACATGAAGCAAGGT 2152
 QY 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTTACACAGAGCTGCAATATGTTGGTCACT 1542
 DB 2153 TAATGCTCTGGATAATCTTGGTCAGACTTCTTACACAGAGCTGCAATATGTTGGTCACT 2212
 QY 1543 ACAAACTGCCGCTACTCTGAGCTATGGGTGATGCTTAAACATATATATCCCTTCAGGG 1602
 DB 2213 ACAAACTGCCGCTACTCTGAGCTATGGGTGATGCTTAAACATATATATCCCTTCAGGG 2272
 QY 1603 CTTTACTGTTTACAGATGGAAATGAAATGTACAGCACTCTTCAAGAGGATATCTC 1662
 DB 2273 CTTTACTGTTTACAGATGGAAATGAAATGTACAGCACTCTTCAAGAGGATATCTC 2332
 QY 1663 ATTAGTAATTCAGAGGCGACAGACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGCTGA 1722
 DB 2333 ATTAGTAATTCAGAGGCGACAGACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGCTGA 2392
 QY 1723 AACTGTAAAAAACTGTACTGTTTACAGAGTGTCAACTGCAGAGACATTAAGAGGGCTCA 1782
 DB 2393 AACTGTAAAAAACTGTACTGTTTACAGAGTGTCAACTGCAGAGACATTAAGAGGGCTCA 2452
 QY 1783 GTCACACACTTCATTTTGCAGCTGGGTATACAGAGTGTCCGCTGGTGAATATCTGCT 1842
 DB 2453 GTCACACACTTCATTTTGCAGCTGGGTATACAGAGTGTCCGCTGGTGAATATCTGCT 2512
 QY 1843 ACAGCATGAGCTGATGTCATGCTTAAAGATAAAGAGGCTTGTACCTTTGCAACAATGC 1902
 DB 2513 ACAGCATGAGCTGATGTCATGCTTAAAGATAAAGAGGCTTGTACCTTTGCAACAATGC 2572
 QY 1903 ATGTTCTTATGGACATTTATGAAGTTGCAGAACTTCTTTGTTAAACATGGAGCAGTAGTTAA 1962

|||||
Db 2573 ATGTTCTTATGGACATATGAAGTTGCGAAGCTTCTTGTTAAACATGGACGAGTAA 2632
|||||
QY 1963 TGTAGCTGATTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATGA 2022
|||||
Db 2633 TGTAGCTGATTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATGA 2692
|||||
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGAGACCTTACCACCAAAAAACAGGATGAAA 2082
|||||
Db 2693 AATTTGCAAACTTCTGCTCCAGCATGGTGAGACCTTACCACCAAAAAACAGGATGAAA 2752
|||||
QY 2083 TACTCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 2142
|||||
Db 2753 TACTCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 2812
|||||
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTCAAGAAGTTGCTTC 2202
|||||
Db 2813 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTCAAGAAGTTGCTTC 2872
|||||
QY 2203 TCTGTATAATGTAATTTGCCCGATACCCAAAGCAGACATTTCAACACCTTTACATTTAGC 2262
|||||
Db 2873 TCTGTATAATGTAATTTGCCCGATACCCAAAGCAGACATTTCAACACCTTTACATTTAGC 2932
|||||
QY 2263 AGCTGGTTATAATTTAGAACTTGCAGAGTATTTGTTTACAAACGGAGCTGATGTGA 2322
|||||
Db 2933 AGCTGGTTATAATTTAGAACTTGCAGAGTATTTGTTTACAAACGGAGCTGATGTGA 2992
|||||
QY 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGAGCATCTTACGGCAGCTAGA 2382
|||||
Db 2993 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGAGCATCTTACGGCAGCTAGA 3052
|||||
QY 2383 TGTAGCAGCTCTACTAATAAAGTATAATGATGTGCAATGCCACGACAAATGGGCTTT 2442
|||||
Db 3053 TGTAGCAGCTCTACTAATAAAGTATAATGATGTGCAATGCCACGACAAATGGGCTTT 3112
|||||
QY 2443 CACACCTTTCCAGAGCAGCCCAAAAGGACCAACAGCTTTTGCTTTGCTAGC 2502
|||||
Db 3113 CACACCTTTCCAGAGCAGCCCAAAAGGACCAACAGCTTTTGCTTTGCTAGC 3172
|||||
QY 2503 CCATGAGCTGACCCGACTCTTAAAAATCAGGAAGCACAACACCTTTAGATTTAGTTTC 2562
|||||
Db 3173 CCATGAGCTGACCCGACTCTTAAAAATCAGGAAGCACAACACCTTTAGATTTAGTTTC 3232
|||||
QY 2563 AGCGATGATGTGAGGCTCTTCTGACAGAGCCATGCCCCCTGCTCTCCCTCTTG 2622
|||||
Db 3233 AGCAGATGATGTGAGGCTCTTCTGACAGAGCCATGCCCCCTGCTCTCCCTCTTG 3292
|||||
QY 2623 TTCAAGCCTCAAGTCTCAATGCTGTGAGAGCCAGAGCCACTGAGATGCTCTCTC 2682
|||||
Db 3293 TTCAAGCCTCAAGTCTCAATGCTGTGAGAGCCAGAGCCACTGAGATGCTCTCTC 3352
|||||
QY 2683 TTCAGGTCATCTAGCCCATCAAGCCTTTCTGACGACGAGCTTTTGACAACCTTATCTGG 2742
|||||
Db 3353 TTCAGGTCATCTAGCCCATCAAGCCTTTCTGACGACGAGCTTTTGACAACCTTATCTGG 3412
|||||
QY 2743 GAGTTTTTCAGAACTCTCTTCAAGTAGTTAGTTCAAGTGAACAGAGGCTGCTTCCAGTTT 2802
|||||
Db 3413 GAGTTTTTCAGAACTCTCTTCAAGTAGTTAGTTCAAGTGAACAGAGGCTGCTTCCAGTTT 3472
|||||
QY 2803 GGAGAAAAAGGAGGTTCCAGGAGTATGATTTAGCATTAAGTCAATTCGTAAGGAATCTTGG 2862
|||||
Db 3473 GGAGAAAAAGGAGGTTCCAGGAGTATGATTTAGCATTAAGTCAATTCGTAAGGAATCTTGG 3532
|||||
QY 2863 ACTTGAGCACCTTAATGGATATATTTGAGAGAAACAGATCACTTTGGATGATTTAGTTGA 2922
|||||
Db 3533 ACTTGAGCACCTTAATGGATATATTTGAGAGAAACAGATCACTTTGGATGATTTAGTTGA 3592
|||||
QY 2923 GATGGGCACAGGAGCTGAAGGAGATTGGAATCAATGCTTTATGACATAGGCACAAACT 2982
|||||
Db 3593 GATGGGCACAGGAGCTGAAGGAGATTGGAATCAATGCTTTATGACATAGGCACAAACT 3652
|||||
QY 2983 AATTAAGGAGTCCGAGAGCTTATCTCCGGACAAAGGCTTTTACCCATATTTAACTTT 3042
|||||

Db 3653 AATTAAGGAGTCCGAGAGCTTATCTCCGGACAAACAGGTCTTTAAACCATATTTAACTTT 3712
QY 3043 GAACACCTCTGCTAGTGGAAACAATTTCTTTATAGATCTGCTCTCTGATGATAAAGATTTCA 3102
|||||
Db 3713 GAACACCTCTGCTAGTGGAAACAATTTCTTTATAGATCTGCTCTCTGATGATAAAGATTTCA 3772
|||||
QY 3103 GTCTGTGGAGGAGAGATGCAAAAGTACAGTTTCGAGAGCACAGATGGAGTTCATGCAGG 3162
|||||
Db 3773 GTCTGTGGAGGAGAGATGCAAAAGTACAGTTTCGAGAGCACAGATGGAGTTCATGCAGG 3832
|||||
QY 3163 TGGAACTTTCAACAGATACAATTTCTCAAGATTCAGAAAGTTTCTAAACAAGAACTATG 3222
|||||
Db 3833 TGGAACTTTCAACAGATACAATTTCTCAAGATTCAGAAAGTTTCTAAACAAGAACTATG 3892
|||||
QY 3223 GGAAGATACACTCACCGGAGAAAAGTTTCTGAAGAAACCAACACCATGCAATGA 3282
|||||
Db 3893 GGAAGATACACTCACCGGAGAAAAGTTTCTGAAGAAACCAACACCATGCAATGA 3952
|||||
QY 3283 ACGAATGCTATTTTCATGGGTCTCTTTTGTGAATCAATTTATCCACAAGGCTTTGATGA 3342
|||||
Db 3953 ACGAATGCTATTTTCATGGGTCTCTTTTGTGAATCAATTTATCCACAAGGCTTTGATGA 4012
|||||
QY 3343 AAGGATGCTACATAGTGGTATGTTTGGAGCTGGCATTTATTTGCTGAAAACCTCTTC 3402
|||||
Db 4013 AAGGATGCTACATAGTGGTATGTTTGGAGCTGGCATTTATTTGCTGAAAACCTCTTC 4072
|||||
QY 3403 CAAAAGCAATCAATATGATATGGAATTTGGAGGAGGTACTGGGTGTCAGTTTCACAAAGA 3462
|||||
Db 4073 CAAAAGCAATCAATATGATATGGAATTTGGAGGAGGTACTGGGTGTCAGTTTCACAAAGA 4132
|||||
QY 3463 CAGATCTTTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGTAACTTTGGGAAAGTC 3522
|||||
Db 4133 CAGATCTTTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGTAACTTTGGGAAAGTC 4192
|||||
QY 3523 TTTCTCGAGTTCAGTGCATGCAATGAAATGGACATTTCTCTCCAGGTCATCAGTCAGTCC 3582
|||||
Db 4193 TTTCTCGAGTTCAGTGCATGCAATGAAATGGACATTTCTCTCCAGGTCATCAGTCAGTCC 4252
|||||
QY 3583 TGGTAGGCCCCAGTGAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGACA 3642
|||||
Db 4253 TGGTAGGCCCCAGTGAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGACA 4312
|||||
QY 3643 GGCTTATCTCTGAGTATTTAAATTTACTTTACAGATTTATGAGGCTGAAGGTATGGTCTGATGG 3702
|||||
Db 4313 GGCTTATCTCTGAGTATTTAAATTTACTTTACAGATTTATGAGGCTGAAGGTATGGTCTGATGG 4372
|||||
QY 3703 ATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGCAGCAGTGG 3762
|||||
Db 4373 ATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGCAGCAGTGG 4432
|||||
QY 3763 CCTCTACGTTTTTACTCCTTTGCTGAAAAAAA 3794
|||||
Db 4433 CCTCTACGTTTTTACTCCTTTGCTGAAAAAAA 4464

RESULT 8

AAF63952

ID AAF63952 standard; DNA; 5002 BP.

XX AAF63952;

AC AAF63952;

DT 05-APR-2001 (first entry)

XX Human tankyrase2 TANK2-LONG coding sequence SEQ ID NO: 132.

DE Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

KW inflammatory disorder; ds.

XX Homo sapiens.

OS Homo sapiens.

XX WO200100849-A1.

XX 04-JAN-2001.

PD 04-JAN-2001.

[illegible]

Db	2763	TACTCTCTTGGATCTTGTTTAAAGATGGAGATACAGATAATCAAGATCTGCTTAGGGGAGA	2822
Qy	2143	TGCAGCTTTGGCTAGATGCTGCCAAGAAGGGTGTGTTACCCAGAGTGAAGAAGTGTGCTTC	2202
Db	2823	TGCAGCTTTGGCTAGATGCTGCCAAGAAGGGTGTGTTACCCAGAGTGAAGAAGTGTGCTTC	2882
Qy	2203	TCCTGATAATGTAAATPGCCGGGATACCCAAGGCAGACATTCACACCTTTACATATTAGC	2262
Db	2883	TCCTGATAATGTAAATPGCCGGGATACCCAAGGCAGACATTCACACCTTTACATATTAGC	2942
Qy	2263	AGCTGGTTATTAATTTAGAAGTTGCAGAGTATTTGTTACACACAGGAGCTGATGTGAA	2322
Db	2943	AGCTGGTTATTAATTTAGAAGTTGCAGAGTATTTGTTACACACAGGAGCTGATGTGAA	3002
Qy	2323	TGCCCAAGACAAAAGGAGGACTTATTCCCTTTACATAATGCAGCATCTTACGGGCATGTAGA	2382
Db	3003	TGCCCAAGACAAAAGGAGGACTTATTCCCTTTACATAATGCAGCATCTTACGGGCATGTAGA	3062
Qy	2383	TGTAGCAGCTCTACTTAATAAGATATAATGCATGTCTAAATGCCACGACAAATGGGCTTT	2442
Db	3063	TGTAGCAGCTCTACTTAATAAGATATAATGCATGTCTAAATGCCACGACAAATGGGCTTT	3122
Qy	2443	CACACCTTTGCACGACGACGCCAAAGGGAGGACACAGACTTTTGTGCTTTGTGCTAGC	2502
Db	3123	CACACCTTTGCACGACGACGCCAAAGGGAGGACACAGACTTTTGTGCTTTGTGCTAGC	3182
Qy	2503	CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACTTTAGATTAGTTTC	2562
Db	3183	CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACTTTAGATTAGTTTC	3242
Qy	2563	AGCGATGATGTCAGCGCTCTTCTGACAGACGCCATGCCCCATCTGCTCTGCCCTCTTG	2622
Db	3243	AGCAGATGATGTCAGCGCTCTTCTGACAGACGCCATGCCCCATCTGCTCTGCCCTCTTG	3302
Qy	2623	TTACAAGCCTCAAGTCTCAATGGTGTGAGAGGCCAGGAGCCACTGCAGATGCTCTCTC	2682
Db	3303	TTACAAGCCTCAAGTCTCAATGGTGTGAGAGGCCAGGAGCCACTGCAGATGCTCTCTC	3362
Qy	2683	TTCAGGTGCCATCTAGCCCATCAAGCCTTCTGCAGCCAGCAGTCTTGACAACTTATCTCGG	2742
Db	3363	TTCAGGTGCCATCTAGCCCATCAAGCCTTCTGCAGCCAGCAGTCTTGACAACTTATCTCGG	3422
Qy	2743	GAGTTTTTCAGAACTGCTCTCAGTAGTTAGTTTCAAGTGGAAACAGAGGTCGCTCCAGTTT	2802
Db	3423	GAGTTTTTCAGAACTGCTCTCAGTAGTTAGTTTCAAGTGGAAACAGAGGTCGCTCCAGTTT	3482
Qy	2803	GGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATACTCAATTCGTAGGAATCTTTGG	2862
Db	3483	GGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATACTCAATTCGTAGGAATCTTTGG	3542
Qy	2863	ACTTGAGCAGCTTAATGGATATATTTGAGAGAAACAGATCACTTTGGATGTATTAGTTGA	2922
Db	3543	ACTTGAGCAGCTTAATGGATATATTTGAGAGAAACAGATCACTTTGGATGTATTAGTTGA	3602
Qy	2923	GATGGGCACAAAGAGCTGAAGGAGATTGGAATCAATGCTTTATGGACATAGGCACAACT	2982
Db	3603	GATGGGCACAAAGAGCTGAAGGAGATTGGAATCAATGCTTTATGGACATAGGCACAACT	3662
Qy	2983	AATTAAGGAGTCGAGAGACTTATCTCCGGACAAAGGCTTTAACCCATATTTTAACTTT	3042
Db	3663	AATTAAGGAGTCGAGAGACTTATCTCCGGACAAAGGCTTTAACCCATATTTTAACTTT	3722
Qy	3043	GMACCTCTGGTAGTGGAACTTCTATAGATCTGTCTCTCATGATAAAGAGTTTCA	3102
Db	3723	GMACCTCTGGTAGTGGAACTTCTATAGATCTGTCTCTCATGATAAAGAGTTTCA	3782
Qy	3103	GTCTGTGGAGGAAGATGCAAGTACAGTTCGAGACACAGAGATGGAGTCAATGCAGG	3162
Db	3783	GTCTGTGGAGGAAGATGCAAGTACAGTTCGAGACACAGAGATGGAGTCAATGCAGG	3842
Qy	3163	TGGAATCTTCACAGATACAATATTTCTCAAGATTCAGAAAGTTTGTACAGAACTATG	3222

Db	3843	TGGAATCTTCAACAGATACAATAATTCTCAAGATTTCAGAAGGTTTGTACACAAGAACTATG	3903
Qy	3223	GGAAAGATACACTCACCGGAGAAAAGAGTCTTCTGAAGAAAACCAACACCATGCGCAATGA	3282
Db	3903	GGAAAGATACACTCACCGGAGAAAAGAGTCTTCTGAAGAAAACCAACACCATGCGCAATGA	3962
Qy	3283	ACGAATGCTATTTTCATGGGTCCTTTTGTGTAATGCAATTTATCCACAAAGGCTTTTGATGA	3342
Db	3963	ACGAATGCTATTTTCATGGGTCCTTTTGTGTAATGCAATTTATCCACAAAGGCTTTTGATGA	4022
Qy	3343	AAGCATGCGCTACATAGGTGGTATGTTTGGAGCTGSCATTTATTTTGTGTAAGAACTCTTC	3402
Db	4023	AAGCATGCGCTACATAGGTGGTATGTTTGGAGCTGSCATTTATTTTGTGTAAGAACTCTTC	4082
Qy	3403	CANAAGCAATCAATATGTAATGGAATTTGGAGGAGGTACTGGGTGCCAGTTCCACAAAGA	3462
Db	4083	CANAAGCAATCAATATGTAATGGAATTTGGAGGAGGTACTGGGTGCCAGTTCCACAAAGA	4142
Qy	3463	CAGATCTTGTAGATTTGCCACAGGAGCTGCTCTTTTGGCCGGGTAACTTTGGGAAAGTC	3522
Db	4143	CAGATCTTGTAGATTTGCCACAGGAGCTGCTCTTTTGGCCGGGTAACTTTGGGAAAGTC	4202
Qy	3523	TTTCCTGCAGTTCAGTGCAATGAAATGGCAATGTCACATCTCTCCAGGTCACTCACTCAGTCAC	3582
Db	4203	TTTCCTGCAGTTCAGTGCAATGAAATGGCAATGTCACATCTCTCCAGGTCACTCACTCAGTCAC	4262
Qy	3583	TGGTAGGCCAGTGTAATGGCCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA	3642
Db	4263	TGGTAGGCCAGTGTAATGGCCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA	4322
Qy	3643	GGCTTATCCTTGAGTATTTAATTACTTACCAGATTTATGAGGCCTGAAGGTATGGTCGATGG	3702
Db	4323	GGCTTATCCTTGAGTATTTAATTACTTACCAGATTTATGAGGCCTGAAGGTATGGTCGATGG	4382
Qy	3703	ATAAATAGTTATTTTAAAGAACTAATTCACGACCACTTAAATATCATCAAGCAGCAGTGG	3762
Db	4383	ATAAATAGTTATTTTAAAGAACTAATTCACGACCACTTAAATATCATCAAGCAGCAGTGG	4442
Qy	3763	CCTCTAGCTTTTACTCCTTTGCTGAAAAAAA 3794	
Db	4443	CCTCTAGCTTTTACTCCTTTGCTGAAAAAAA 4474	
RESULT 9			
ID	AAD02578		
XX	AAD02578 standard; DNA; 4512 BP.		
AC	AAD02578;		
XX			
DT	02-MAY-2001 (first entry)		
XX			
DE	Human tankyrase homolog protein (THP) DNA.		
XX			
KW	Human; tankyrase homolog protein; THP; gene therapy; cancer;		
KW	tumour; basal cell carcinoma; therapy; genetic mapping;		
KW	cytostatic; ds.		
XX			
OS	Homo sapiens.		
XX			
FH	Key		
FT	CDS		
FT	Location/Qualifiers		
FT	23..3523		
FT	/*tag= a		
FT	/product= "Human tankyrase homolog protein (THP)"		
FT	/note= "This region is specifically claimed in		
FT	claim 4 as SEQ ID NO:4"		
XX			
PN	WO200104326-A1.		
XX			
PD	18-JAN-2001.		
XX			
PF	03-JUL-2000; 2000WO-EP06609.		
XX			
PR	09-JUL-1999; 99US-0350982.		
XX			

XX (PHAA) PHARMACIA & UPJOHN SPA.
PA Berthelsen J, Toma S, Isacchi A;
PI
XX
XX
DR WPI: 2001-168422/17.
DR P-PSDB; AAY72589.
XX
PT New tankyrase homolog protein (THP) polynucleotide and polypeptide
PT useful in gene therapy, diagnosis and treatment or prevention of
PT unregulated cell growth, such as cancer or tumor cell growth -
XX
PS Claim 1; Page 41-42; 60pp; English.
XX
CC The present sequence is a DNA coding for human tankyrase homolog
CC protein (THP). The THP polypeptides and polynucleotides of the
CC invention are useful in gene therapy and for treating or preventing
CC unregulated cell growth such as cancer or tumour (e.g. basal cell
CC carcinoma). The nucleic acid molecules of the invention and their
CC fragments are useful for restriction fragment length polymorphism
CC (RFLP) associated with certain disorders, as well as for genetic
CC mapping. Antisense oligonucleotides, or fragments of nucleic acid
CC encoding THP are useful as diagnostic tools for probing the
CC expression of Thp gene in various tissues. THP can be used as
CC antigens for raising antibodies against them and in assays for
CC identifying compounds that modulate their activity. They are used in
CC the manufacture of a medicament directed towards cancers or tumours.
CC THP are also useful for screening compounds in a variety of drugs
CC screening techniques and as a research tool for identification,
CC characterisation and purification of interacting, regulatory proteins.
XX
SQ Sequence 4512 BP; 1356 A; 878 C; 1002 G; 1268 T; 8 other;

Query Match 89.1%; Score 3382.8; DB 22; Length 4512;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3381; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 403 AGGTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAAGTCCCAAGC 462
DB |||||
DB 220 AGGTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAAGTCCCAAGC 279
QY 463 ACCTGATGATGGGGCTTATTCCTCTTCATAATGATGCTCTTTGGTCATGCTGAAGT 522
DB |||||
DB 280 ACCTGATGATGGGGCTTATTCCTCTTCATAATGATGCTCTTTGGTCATGCTGAAGT 339
QY 523 AGTCAATCTCTTTTGGACATGGTGCAGACCCCAATGCTCGAGATAAATGGAAATTATAC 582
DB |||||
DB 340 AGTCAATCTCTTTTGGACATGGTGCAGACCCCAATGCTCGAGATAAATGGAAATTATAC 399
QY 583 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
DB |||||
DB 400 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 459
QY 643 TGGAGCTGAGCCCAACCATCGGAATACAGATGGAAGACAGCAATTTGATGTTTACAGATCC 702
DB |||||
DB 460 TGGAGCTGAGCCCAACCATCGGAATACAGATGGAAGACAGCAATTTGATGTTTACAGATCC 519
QY 703 ATCTGCCAAAGCAGTCTTACTGCTGAATATAAGAAAGATGAAGTCTTAGAAAGTGCAG 762
DB |||||
DB 520 ATCTGCCAAAGCAGTCTTACTGCTGAATATAAGAAAGATGAAGTCTTAGAAAGTGCAG 579
QY 763 GAGTGGCAATGAAGAAAAATGATGCTCTTACTCACACCAATTTAAATGTCAACTGCCACGC 822
DB |||||
DB 580 GAGTGGCAATGAAGAAAAATGATGCTCTTACTCACACCAATTTAAATGTCAACTGCCACGC 639
QY 823 AAGTGTATGCCAGAAAGTCACTCCATTACATTTGGCAGCAGGATATAACAGAGATAAGAT 882
DB |||||
DB 640 AAGTGTATGCCAGAAAGTCACTCCATTACATTTGGCAGCAGGATATAACAGAGATAAGAT 699
QY 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTTAAAGATAAAGGTCATCGGT 942
DB |||||
DB 700 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTTAAAGATAAAGGTCATCGGT 759

QY 943 ACCATTACACAATGCCCTGTTCTTTATGTCATTATGAAGTAACCTGAACCTTTTGGTCAAGCA 1002
DB |||||
DB 760 ACCATTACACAATGCCCTGTTCTTTATGTCATTATGAAGTAACCTTTTGGTCAAGCA 819
QY 1003 TGGTGCTCTGTTAAATGCAATGGACTTTGTGCAATTCACCTCTTTCATGAGGAGCTTC 1062
DB |||||
DB 820 TGGTGCTCTGTTAAATGCAATGGACTTTGTGCAATTCACCTCTTTCATGAGGAGCTTC 879
QY 1063 TAGAACACAGGTTGAAGTATGTTCTCTTCTTAAAGTATGTTGGTCAGACCCACACACTGCT 1122
DB |||||
DB 880 TAGAACACAGGTTGAAGTATGTTCTCTTCTTAAAGTATGTTGGTCAGACCCACACACTGCT 939
QY 1123 CAATTGTCACAATAAAAGTCTATAGACTTGGCTCCACACACACAGTTAAAGAAAGATT 1182
DB |||||
DB 940 CAATTGTCACAATAAAAGTCTATAGACTTGGCTCCACACACAGTTAAAGAAAGATT 999
QY 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGAGAACTGATGTTACTCG 1242
DB |||||
DB 1000 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGAGAACTGATGTTACTCG 1059
QY 1243 AATCAAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGCATCCTCAACACATGAAC 1302
DB |||||
DB 1060 AATCAAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGCATCCTCAACACATGAAC 1119
QY 1303 AGCATTGCATTGCTGCTGCTATCCATATCCCAAGAAAGCAAAATATCTGAACCTGTT 1362
DB |||||
DB 1120 AGCANTGCATTGCTGCTGCTATCCATATCCCAAGAAAGCAAAATATGTAACCTGTT 1179
QY 1363 GCTAAGAAAGAGGCAACATCAATGAAAGACCTTAAAGAAATTTCTTGACTCTCTGCACT 1422
DB |||||
DB 1180 GCTAAGAAAGAGGCAACATCAATGAAAGACCTTAAAGAAATTTCTTGACTCTCTGCACT 1239
QY 1423 GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGTGAACATCAAGCAAAAGST 1482
DB |||||
DB 1240 GGCATCTGAGAAAGCTCATATGATRTGTTGAAGTAGTGTGAACATCAAGCAAAAGST 1299
QY 1483 TAATGCTCTGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGTCATCT 1542
DB |||||
DB 1300 TAATGCTCTGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGTCATCT 1359
QY 1543 ACAAACTGCCGCTTACTCTGAGCTATGGGTGTGATCCTTAACATATATATCCTTTCAGG 1602
DB |||||
DB 1360 ACAAACTGCCGCTTACTCTGAGCTATGGGTGTGATCCTTAACATATATATCCTTTCAGG 1419
QY 1603 CTTTACTGCTTTACAGATGGCAATGAAATGTACAGCAACTCTCTCAAGAGGATATCTC 1662
DB |||||
DB 1420 CTTTACTGCTTTACAGATGGCAATGAAATGTACAGCAACTCTCTCAAGAGGATATCTC 1479
QY 1663 ATTAGGTAATTCAGAGGACAGACAGCAATTTGCTGGAAGCTGCAAGAGCTGGAGATGTCGA 1722
DB |||||
DB 1480 ATTAGGTAATTCAGAGGACAGACAGCAATTTGCTGGAAGCTGCAAGAGCTGGAGATGTCGA 1539
QY 1723 AACTGTAAAAAACTGTGACTGTTGAGAGTGTCAACTGCAGAGACATTTGAAGGGCTGCA 1782
DB |||||
DB 1540 AACTGTAAAAAACTGTGACTGTTGAGAGTGTCAACTGCAGAGACATTTGAAGGGCTGCA 1599
QY 1783 GTCTACACCACTTCATTTTGCAGCTGGGTAAACAGAGTGTGCTGCTGGAATATCTGCT 1842
DB |||||
DB 1600 GTCTACACCACTTCATTTTGCAGCTGGGTAAACAGAGTGTGCTGCTGGAATATCTGCT 1659
QY 1843 ACAGCATGGAGCTGATGTGCATCTAAAAGATAAGGAGGCTTTGTACCTTTTGCACAATGC 1902
DB |||||
DB 1660 ACAGCATGGAGCTGATGTGCATCTAAAAGATAAAGRRGSCCTTGTACCTTTTGCACAATGC 1719
QY 1903 ATGTTCTTATGGACATATGAAGTTGCAGAACTCTTGTAAACATGGAGCAGTAGTTAA 1962
DB |||||
DB 1720 ATGTTCTTATGGACATATGAAGTTGCAGAACTCTTGTAAACATGGAGCAGTAGTTAA 1779
QY 1963 TGTAGCTGATTTATGAAATTTACACCTTTTACATGAAGCAGCAGCAAAAGCAAAATATGA 2022
DB |||||
DB 1780 TGTAGCTGATTTATGAAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGCAAAATATGA 1839
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTTACCACAAAAAAGAGGATGGAAA 2082

Db 1840 AATTTGGAACTTCTGCTCCAGCATGTCGAGACCTTACAAAAAACAAGGATGAAA 1899
Qy 2083 TACTCCTTTGGATCTTGTAAAGATGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 2142
Db 1900 TACTCCTTTGGATCTTGTAAAGATGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 1959
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTACGAGAGTGAAGAGTTGCTTC 2202
Db 1960 TGCAGCTTTGCTAGATGCTGCCAAGAGAGTTGTTTACGAGAGTGAAGAGTTGCTTC 2019
Qy 2203 TCCTGATAATGTAATTTGCCGCGATACCCAGGAGAGTTCACACCTTTTACATTTAGC 2262
Db 2020 TCCTGATAATGTAATTTGCCGCGATACCCAGGAGAGTTCACACCTTTTACATTTAGC 2079
Qy 2263 AGCTGGTTTATTAATTTAGAGTTGAGAGTATTTGTTTACACACGAGCTGATGTGA 2322
Db 2080 AGCTGGTTTATTAATTTAGAGTTGAGAGTATTTGTTTACACACGAGCTGATGTGA 2139
Qy 2323 TGCCCAAGACAAAGAGGACTTATTCCTTTTACATAATGACAGCATCTTACGGCATGTGA 2382
Db 2140 TGCCCAAGACAAAGAGGACTTATTCCTTTTACATAATGACAGCATCTTACGGCATGTGA 2199
Qy 2383 TGTAGAGCTCTACTAATAAAGTATAATGATGTGTCATGTCATGTCAGGAGCTTTGTCGTAGC 2502
Db 2200 TGTAGAGCTCTACTAATAAAGTATAATGATGTGTCATGTCATGTCAGGAGCTTTGTCGTAGC 2259
Qy 2443 CACACCTTTGACAGNAGAGCCCAAGAGGAGGAGACAGACAGCTTTGTCGTAGC 2502
Db 2260 CACACCTTTGACAGNAGAGCCCAAGAGGAGGAGACAGACAGCTTTGTCGTAGC 2319
Qy 2503 CCATGGAGCTGACCGGACTCTTAAATAATCAGGAGGAGCAACACCTTTAGATTTAGTTTC 2562
Db 2320 CCATGGAGCTGACCGGACTCTTAAATAATCAGGAGGAGCAACACCTTTAGATTTAGTTTC 2379
Qy 2563 AGCGGATGATGTACGCGCTCTTCTGACAGCAGCCATGCCCCATCTGCTCTGCGCCCTTG 2622
Db 2380 AGCGGATGATGTACGCGCTCTTCTGACAGCAGCCATGCCCCATCTGCTCTGCGCCCTTG 2439
Qy 2623 TTACAGACCTCAAGTGTCTCAATGGTGTGAGAGAGCCAGGAGCCACTGCAGATGCTCTC 2682
Db 2440 TTACAGACCTCAAGTGTCTCAATGGTGTGAGAGAGCCAGGAGCCACTGCAGATGCTCTC 2499
Qy 2683 TTACAGTCCATCTAGCCCATCAAGCTTTCTGACGCCAGCAGTCTGACAACTTTATCTGG 2742
Db 2500 TTACAGTCCATCTAGCCCATCAAGCTTTCTGACGCCAGCAGTCTTACAACTTTATCTGG 2559
Qy 2743 GAGTTTTTCAGAACTGTCTTCTCAGTAGTTAGTTCAAGTGGAAACAGAGGCTGCTCCAGTTT 2802
Db 2560 GAGTTTTTCAGAACTGTCTTCTCAGTAGTTAGTTCAAGTGGAAACAGAGGCTGCTCCAGTTT 2619
Qy 2803 GGAGAAAGAGGAGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2620 GGAGAAAGAGGAGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2679
Qy 2863 ACTTGAGCACCTTAATGGATATATTTGAGAGAGACAGATCACTTTGGATGATTAGTTGA 2922
Db 2680 ACTTGAGCACCTTAATGGATATATTTGAGAGAGACAGATCACTTTGGATGATTAGTTGA 2739
Qy 2923 GATGGGGCACAGGAGCTGAAGGAGATTGGAATCAATGCTTATGACATAGGACACAACT 2982
Db 2740 GATGGGGCACAGGAGCTGAAGGAGATTGGAATCAATGCTTATGACATAGGACACAACT 2799
Qy 2983 AATTAAGGAGTCGAGAGACTTATCTCCGACACAAAGGCTTAACCCATATTTAACTTT 3042
Db 2800 AATTAAGGAGTCGAGAGACTTATCTCCGACACAAAGGCTTAACCCATATTTAACTTT 2859
Qy 3043 GAACACCTCTGCTAGTGAACAATTTCTTATAGATCTGCTCTGATGATAAAGAGTTTCA 3102
Db 2860 GAACACCTCTGCTAGTGAACAATTTCTTATAGATCTGCTCTGATGATAAAGAGTTTCA 2919
Qy 3103 GTCTGTGGAGGAGAGATGCAAGTACAGTTTCGAGAGCACAGAGATGGAGTCAATGCAGG 3162

Db 2920 GTCTGTGGAGGAGAGATGCAAAAGTACAGTTTCGAGAGCACAGAGATGGAGGTCATCCAGG 2979
Qy 3163 TGGAAATCTTCAACAGATACAATATTTCTCAAGATTCAGAGGTTTGTAAACAAGAACTATG 3222
Db 2980 TGGAAATCTTCAACAGATACAATATTTCTCAAGATTCAGAGGTTTGTAAACAAGAACTATG 3039
Qy 3223 GGAAGATACACTCACCGGAGAGAAAGTTTCTGAAAGAAAACCAACCATGCCAATGA 3282
Db 3040 GGAAGATACACTCACCGGAGAGAAAGTTTCTGAAAGAAAACCAACCATGCCAATGA 3099
Qy 3283 ACGAATGCTATTTTCATGGGTCCTTTTGTGATGCAATATTCACAAAGGCTTTGATGA 3342
Db 3100 ACGAATGCTATTTTCATGGGTCCTTTTGTGATGCAATATTCACAAAGGCTTTGATGA 3159
Qy 3343 AAGGCATGCTACATAGGTGGTATGTTTGGAGCTGCAATTTATTTGCTGAAAACCTTTC 3402
Db 3160 AAGGCATGCTACATAGGTGGTATGTTTGGAGCTGCAATTTATTTGCTGAAAACCTTTC 3219
Qy 3403 CAAAAGCAATCAATATGATATGGAATTTGGAGAGGTACTGGGTGTCAGTTTCAAAAGA 3462
Db 3220 CAAAAGCAATCAATATGATATGGAATTTGGAGAGGTACTGGGTGTCAGTTTCAAAAGA 3279
Qy 3463 CAGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGGCGGTAACCTTTGGGAAGTC 3522
Db 3280 CAGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGGCGGTAACCTTTGGGAAGTC 3339
Qy 3523 TTTCTGTCAGTTTCAGTGAATGAAATGGCACAATTTCTCTCCAGGTCATCACTCAGTCAC 3582
Db 3340 TTTCTGTCAGTTTCAGTGAATGAAATGGCACAATTTCTCTCCAGGTCATCACTCAGTCAC 3399
Qy 3583 TGTAGGCCCAAGTGAATGGCTAGCATTAAGTGAATATGTTTATTTACAGAGGAGAAACA 3642
Db 3400 TGTAGGCCCAAGTGAATGGCTAGCATTAAGTGAATATGTTTATTTACAGAGGAGAAACA 3459
Qy 3643 GGCTTATCTGAGATTTAATTAATTTACAGATTTAGAGCCCTGAAGGATGTTGTCATGG 3702
Db 3460 GGCTTATCTGAGATTTAATTAATTTACAGATTTAGAGCCCTGAAGGATGTTGTCATGG 3519
Qy 3703 ATAAATAGTTTATTTTAAAGAACTAATTTCCACTGAACCTAAATCATCAAGACAGCAGTGG 3762
Db 3520 ATAAATAGTTTATTTTAAAGAACTAATTTCCACTGAACCTAAATCATCAAGACAGCAGTGG 3579
Qy 3763 CCTCTAGCTTTTACTCCTTTTGTGCTGAAAAAAA 3794
Db 3580 CCTCTAGCTTTTACTCCTTTTGTGCTGAAAAAAA 3611

RESULT 10
AA25366
ID AAX25366 standard; cDNA; 3400 BP.
XX
AC AAX25366;
XX
DT 19-JUL-1999 (first entry)
XX
DE Human Grb7 effector 2.2412 cDNA.
XX
KW Grb7 effector; 2.2412 protein; human; signal transduction;
tumour marker; breast cancer; prostate cancer; prognosis;
diagnosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..3225
FT /*Lag= a
XX
XX WO9915647-A1.
XX
XX 01-APR-1999.
XX
XX 23-SEP-1998; 98WO-AU00795.
XX

PR 23-SEP-1997; 97AU-0009388.
 PA (GARV-) GARVAN INST MEDICAL RES.
 XX Daily RJ, Sutherland RL;
 XX WPI; 1999-254707/21.
 DR P-PSDB; AAY05734.
 XX
 PT New candidate effector for the Grb7 family of signaling proteins,
 PT and specific antibody, useful for detection and treatment of cancer
 XX
 PS Claim 4; Fig 1; 24pp; English.
 XX
 CC This is the nucleotide sequence of an isolated polynucleotide
 CC molecule encoding a candidate effector protein, termed 2.2412 (see
 CC AAY05734), for the Grb7 family of signalling proteins. A partial
 CC clone was isolated from a human liver cDNA library using a yeast
 CC two-hybrid system assay with Grb14 as bait. This clone was used
 CC to screen a human placental cDNA library that provided the
 CC c-terminal end of the 2.2412 sequence. A 5' sequence for the
 CC clone, including the initiation codon, was not identified.
 CC Analysis of the sequence revealed significant homology to a large
 CC number of proteins containing ankyrin-like repeats. The gene
 CC was localised to between chromosome 10q23.2 and proximal 10q23.32.
 CC Deletions in the 10q22-25 region have been detected in human breast,
 CC prostate, renal, small cell lung and endometrial carcinomas,
 CC glioblastoma multiforme, melanoma and meningiomas. Detection of
 CC the protein encoded by the 2.2412 cDNA in a sample should provide a
 CC useful tumour marker and/or prognostic indicator for certain human
 CC cancers, in particular breast cancer and prostate cancer.
 CC Antagonism of the interaction between Grb7 family members and the
 CC encoded protein should provide a novel treatment strategy for human
 CC diseases exhibiting aberrant receptor tyrosine kinase signalling,
 CC such as cancer. Oligonucleotide probes can be used in methods of
 CC detecting the presence of 2.2412 mRNA in a sample.
 XX
 SQ Sequence 3400 BP; 1059 A; 565 G; 760 G; 916 T; 0 other;

Query Match 86.8%; Score 3297; DB 20; Length 3400;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 3303; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Db	421	CATGGACGTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT	480
Qy	962	TCTTATGTCATTATGAAGTAACCTGAACTTTTGGTCAAGCATGCTGCTGTAATGCA	1021
Db	481	TCTTATGTCATTATGAAGTAACCTGAACTTTTGGTCAAGCATGCTGCTGTAATGCA	540
Qy	1022	ATGACTTGTGGCAATTCACCTCTTTCATGAGCAGCTTCTTAAGAACAGCGTTGAAGTA	1081
Db	541	ATGACTTGTGGCAATTCACCTCTTTCATGAGCAGCTTCTTAAGAACAGCGTTGAAGTA	600
Qy	1082	TGTTCTCTCTTCTTAAAGTTATGGTGCAGACCCAACTGCTCAATTTGTCACAATAAAAGT	1141
Db	601	TGTTCTCTCTTCTTAAAGTTATGGTGCAGACCCAACTGCTCAATTTGTCACAATAAAAGT	660
Qy	1142	GCTATAGACTTGGCTCCACACACAGCTTAAAGAACAGATTAGCATATGAATTTAAAGGC	1201
Db	661	GCTATAGACTTGGCTCCACACACAGCTTAAAGAACAGATTAGCATATGAATTTAAAGGC	720
Qy	1202	CACCTCGTGTGCGCAAGCTGCACGAGAGAGCTGATCTTACTCGAATCAAAAACATCTCTCT	1261
Db	721	CACCTCGTGTGCGCAAGCTGCACGAGAGAGCTGATCTTACTCGAATCAAAAACATCTCTCT	780
Qy	1262	CTGAAATGTTGAATTTCAAGCATCTTCAACACATGAAACAGCATTTGCTGCTGT	1321
Db	781	CTGAAATGTTGAATTTCAAGCATCTTCAACACATGAAACAGCATTTGCTGCTGT	840
Qy	1322	GCATCTCCATATCCCAAGAACCAATATGTGAACCTGTTGCTAAGAAAGAGCAAAAC	1381
Db	841	GCATCTCCATATCCCAAGAACCAATATGTGAACCTGTTGCTAAGAAAGAGCAAAAC	900
Qy	1382	ATCAATGAAAGACTTAAAGAAATCTTGACTCTCTGACCTGGCATCTGAGAAGCTCAT	1441
Db	901	ATCAATGAAAGACTTAAAGAAATCTTGACTCTCTGACCTGGCATCTGAGAAGCTCAT	960
Qy	1442	AATCATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGGTTAATGCTGCGATATCTT	1501
Db	961	AATCATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGGTTAATGCTGCGATATCTT	1020
Qy	1502	GGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATCTACAAACCTGCCGCCCTACTC	1561
Db	1021	GGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATCTACAAACCTGCCGCCCTACTC	1080
Qy	1562	CTGAGCTATGGTGTGATCCTAACATTTATCCCTTTCAGGGCTTTTACTGCTTTACAGATG	1621
Db	1081	CTGAGCTATGGTGTGATCCTAACATTTATCCCTTTCAGGGCTTTTACTGCTTTACAGATG	1140
Qy	1622	GGAATGAAATGTACAGCAACTCCCTCCAGAGGGTATCTCATTTAGGTAATTCAGAGGCA	1681
Db	1141	GGAATGAAATGTACAGCAACTCCCTCCAGAGGGTATCTCATTTAGGTAATTCAGAGGCA	1200
Qy	1682	GACAGCAATTTGCTGGAAGCTGCAAGGCTGGAGATGTGCAAACTGTAAAAAACTGTGT	1741
Db	1201	GACAGCAATTTGCTGGAAGCTGCAAGGCTGGAGATGTGCAAACTGTAAAAAACTGTGT	1260
Qy	1742	ACTGTTCCAGAGTGTCAACTGCAGAGCAATGAAGGGCGTCAGCTACACACTTCATTTT	1801
Db	1261	ACTGTTCCAGAGTGTCAACTGCAGAGCAATGAAGGGCGTCAGCTACACACTTCATTTT	1320
Qy	1802	GCAGCTGGGTATACAGAGTCCGCTGGTGAATATCTGCTACAGCATGAGCTGATGTG	1861
Db	1321	GCAGCTGGGTATACAGAGTCCGCTGGTGAATATCTGCTACAGCATGAGCTGATGTG	1380
Qy	1862	CATGCTAAAGATAAAGGAGGCTTTGACCTTTGCACAATGCATGTTCTTATGGACATTAT	1921
Db	1381	CATGCTAAAGATAAAGGAGGCTTTGACCTTTGCACAATGCATGTTCTTATGGACATTAT	1440
Qy	1922	GAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTAAATGCTAGCTGATTTATGGAAA	1981
Db	1441	GAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTAAATGCTAGCTGATTTATGGAAA	1500
Qy	1982	TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTCTGCTC	2041
Db	1501	TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTCTGCTC	1560

QY 2042 CAGCATGGTCAGACCCCTACCAAAAAAACAAGGATCGAAATACCTCTTTGGATCTGTGTT 2101
Db 1561 CAGCATGGTCAGACCCCTACCAAAAAAACAAGGATCGAAATACCTCTTTGGATCTGTGTT 1620
QY 2102 AAAGATGGAGATACAGATATTCAGATCTCGCTTAGGGAGATGACAGCTTTGCTAGATGCT 2161
Db 1621 AAAGATGGAGATACAGATATTCAGATCTCGCTTAGGGAGATGACAGCTTTGCTAGATGCT 1680
QY 2162 GCCAAGAGGGTGTCTTAGCCAGAGTGAAGAAGTTGCTCTCTGATGATGTAATGTAATGTC 2221
Db 1681 GCCAAGAGGGTGTCTTAGCCAGAGTGAAGAAGTTGCTCTCTGATGATGTAATGTAATGTC 1740
QY 2222 CGCGATACCCAAAGCAGACATTCACACCTTTACATTTAGCAGCTGCTTATAATATTTA 2281
Db 1741 CGCGATACCCAAAGCAGACATTCACACCTTTACATTTAGCAGCTGCTTATAATATTTA 1800
QY 2282 GAAATGTCAGAGTATTTGTTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGAGGGA 2341
Db 1801 GAAATGTCAGAGTATTTGTTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGAGGGA 1860
QY 2342 CTTATTCCTTTACATATGACAGATCTTACGGGATGTAGATGTAGCAGCTCTACTAATA 2401
Db 1861 CTTATTCCTTTACATATGACAGATCTTACGGGATGTAGATGTAGCAGCTCTACTAATA 1920
QY 2402 AAGTATAATGATGTCTCAATGCCAGGACAAATGGGCTTTACACACCTTTGSCACGAAGCA 2461
Db 1921 AAGTATAATGATGTCTCAATGCCAGGACAAATGGGCTTTACACACCTTTGSCACGAAGCA 1980
QY 2462 GCCCAAAAGGAGCAACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACT 2521
Db 1981 GCCCAAAAGGAGCAACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACT 2040
QY 2522 CTTAAATCAGGAGGACAAACACCTTTAGATTTAGTTTTCAGCGGATGTAGCGGCT 2581
Db 2041 CTTAAATCAGGAGGACAAACACCTTTAGATTTAGTTTTCAGCGAGATGTAGCGGCT 2100
QY 2582 CTTCTGACAGACGACATGCCCATCTCTCTGCTGCTTACAGACCTTCAAGTGTCTC 2641
Db 2101 CTTCTGACAGACGACATGCCCATCTCTCTGCTGCTTACAGACCTTCAAGTGTCTC 2160
QY 2642 AATGTTGTGAGAGCCCGAGGAGCCACTGCAGATGCTCTCTCTCAGGTCCATCTAGCCCA 2701
Db 2161 AATGTTGTGAGAGCCCGAGGAGCCACTGCAGATGCTCTCTCTCAGGTCCATCTAGCCCA 2220
QY 2702 TCAAGCCTTTCTGACGACGAGCTCTTACACACTTATCTGGAGTTTTCAGAACTGTCT 2761
Db 2221 TCAAGCCTTTCTGACGACGAGCTCTTACACACTTATCTGGAGTTTTCAGAACTGTCT 2280
QY 2762 TCAGTAGTTAGTTCAAGTGGACAGAGGCTGCTTCCAGTTTGGAAAAAGAGGTTCCA 2821
Db 2281 TCAGTAGTTAGTTCAAGTGGACAGAGGCTGCTTCCAGTTTGGAAAAAGAGGTTCCA 2340
QY 2822 GGAGTAGATTTTAGCATACTCAATTCGTAGGAATCTTGGACTTTGACACCTTAATGGAT 2881
Db 2341 GGAGTAGATTTTAGCATACTCAATTCGTAGGAATCTTGGACTTTGACACCTTAATGGAT 2400
QY 2882 ATATTTGAGAGAGACAGATCACTTTGGATGTATTTAGTGTAGATGGGCAAGAGGCTG 2941
Db 2401 ATATTTGAGAGAGACAGATCACTTTGGATGTATTTAGTGTAGATGGGCAAGAGGCTG 2460
QY 2942 AAGGAGATTGGAACTCAATGCTTATGGATAGGCAACAACTAAATTAAGGAGTTCGAGAGA 3001
Db 2461 AAGGAGATTGGAACTCAATGCTTATGGATAGGCAACAACTAAATTAAGGAGTTCGAGAGA 2520
QY 3002 CTTATCTCGGACACAAAGGCTTTAACCCATATTTAACTTTTGAACACCTCTGGTAGTGA 3061
Db 2521 CTTATCTCGGACACAAAGGCTTTAACCCATATTTAACTTTTGAACACCTCTGGTAGTGA 2580
QY 3062 ACAATCTTTATAGATCTGTCTCTCATGATGAAGAGTTTCAGTCTGTGGAGCAAGAGATG 3121
Db 2581 ACAATCTTTATAGATCTGTCTCTCATGATGAAGAGTTTCAGTCTGTGGAGCAAGAGATG 2640

QY 3122 CAAAGTACAGTTTCGAGAGCACAGAGATGAGGTCATGCAGGTGGAATCTTCAACAGATAC 3181
Db 2641 CAAAGTACAGTTTCGAGAGCACAGAGATGAGGTCATGCAGGTGGAATCTTCAACAGATAC 2700
QY 3182 AATATTCTCAAGATTCAGAAGTTTGTACAAGAACTATGGGAAAGATACACTCACCGG 3241
Db 2701 AATATTCTCAAGATTCAGAAGTTTGTACAAGAACTATGGGAAAGATACACTCACCGG 2760
QY 3242 AGAAAGAAGTTTCTCAAGAAACACCAACCATGCCAATGAACGAATGCTATTTCATGG 3301
Db 2761 AGAAAGAAGTTTCTCAAGAAACACCAACCATGCCAATGAACGAATGCTATTTCATGG 2820
QY 3302 TCTCCTTTTCTGAATTCGAATTCACAAAGGCTTTGATGAAGGCGATCGGTACATAGT 3361
Db 2821 TCTCCTTTTCTGAATTCGAATTCACAAAGGCTTTGATGAAGGCGATCGGTACATAGT 2880
QY 3362 GGTATGTTGGAGCTGGCATTTATTTTGTGTAAGACTCTTCCAAAGCAATCAATATGTA 3421
Db 2881 GGTATGTTGGAGCTGGCATTTATTTTGTGTAAGACTCTTCCAAAGCAATCAATATGTA 2940
QY 3422 TATGGAATTCGAGSAGTACTGGGTGCTCCAGTTTCACAAACAGACAGATCTTGTACATTTGC 3481
Db 2941 TATGGAATTCGAGSAGTACTGGGTGCTCCAGTTTCACAAACAGACAGATCTTGTACATTTGC 3000
QY 3482 CACAGGCGAGCTGCTCTTTTGGCGGGTAACTTGGGAAAGTCTTCTCGAGTTTCACTGCA 3541
Db 3001 CACAGGCGAGCTGCTCTTTTGGCGGGTAACTTGGGAAAGTCTTCTCGAGTTTCACTGCA 3060
QY 3542 ATGAAATGTCACATTTCTCTCCAGGTCTACTCTAGTCTACTGCTAGGCCCCAGTGTAAAT 3601
Db 3061 ATGAAATGTCACATTTCTCTCCAGGTCTACTCTAGTCTACTGCTAGGCCCCAGTGTAAAT 3120
QY 3602 GGCTAGCATTTAGCTGAATATGTTTATTTACAGAGGAGAACAGGCTTATCTCGAGTATTTA 3661
Db 3121 GGCTAGCATTTAGCTGAATATGTTTATTTACAGAGGAGAACAGGCTTATCTCGAGTATTTA 3180
QY 3662 ATTACTTACCAGATTTATGAGGCTCTGAAGGTATGTCGATGGATAAATAGTATTTTAA 3721
Db 3181 ATTACTTACCAGATTTATGAGGCTCTGAAGGTATGTCGATGGATAAATAGTATTTTAA 3240
QY 3722 AACTAATTCACACTCAACCTAAATCATCAAGCAGCAGTGGCCTCTACGTTTACTCCTT 3781
Db 3241 AACTAATTCACACTCAACCTAAATCATCAAGCAGCAGTGGCCTCTACGTTTACTCCTT 3300
QY 3782 TGCTGAAAAA 3794
Db 3301 TGCTGAAAAA 3313

RESULT 11

AAF63837

ID AAF63837 standard; DNA; 3508 BP.

XX AAF63837;

XX 05-APR-2001 (first entry)

DT Human tankyrase2 related coding sequence SEQ ID NO: 1.

XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

XX inflammatory disorder; ds.

XX Homo sapiens.

OS WO200100849-A1.

PN 04-JAN-2001.

PD 28-JUN-2000; 2000WO-US17827.

PF 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX WPI: 2001-102896/11.
DR P-PSDB: AAB66278.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders
XX
PS Disclosure; Page 105-109; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
SQ Sequence 3508 BP; 1046 A; 710 C; 843 G; 909 T; 0 other;

Query Match 86.8%; Score 3296.2; DB 22; Length 3508;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3298; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 403 AGTTTGGGGGGAAGACGTAGTGAATATTTGCTTCAGATGGGCAAGTCTCCAGC 462
DB 208 AGTTTGGGGGGAAGACGTAGTGAATATTTGCTTCAGATGGGCAAGTCTCCAGC 267
QY 463 ACCTGATGATGGGGGCTTATCCTCTTCATATGATGCTCTTTTGGTCATGCTGAAGT 522
DB 268 ACCTGATGATGGGGGCTTATCCTCTTCATATGATGCTCTTTTGGTCATGCTGAAGT 327
QY 523 AGTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGAATATAC 582
DB 328 AGTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGAATATAC 387
QY 583 TCCTCTCCATGAAGCTGCAATTAAGGAAGATTTGATGTTTTCATGCTGCTTTACAGCA 642
DB 388 TCCTCTCCATGAAGCTGCAATTAAGGAAGATTTGATGTTTTCATGCTGCTTTACAGCA 447
QY 643 TGAGCTGAGCCCAACCATCGGAATACAGATGGAAGGACAGCATTTGATTTAGCAGATCC 702
DB 448 TGAGCTGAGCCCAACCATCGGAATACAGATGGAAGGACAGCATTTGATTTAGCAGATCC 507
QY 703 ATCTGCCAAAGCAGTCTTACTGGTGAATATTAAGAAAGATGAACCTTAGAAGTGCCAG 762
DB 508 ATCTGCCAAAGCAGTCTTACTGGTGAATATTAAGAAAGATGAACCTTAGAAGTGCCAG 567
QY 763 GAGTGGCAATGAAGAAAAATGATGCTCTACTCACCATTTAAATGTCAACTGCCAGGC 822
DB 568 GAGTGGCAATGAAGAAAAATGATGCTCTACTCACCATTTAAATGTCAACTGCCAGGC 627
QY 823 AGTGTATGCGAAGTCAACTCCATTTACATTTGGCAGAGATATACAGATTAAGAT 882
DB 628 AGTGTATGCGAAGTCAACTCCATTTACATTTGGCAGAGATATTAAGATTAAGAT 687
QY 883 TGTACAGCTGTTTACTGCAACATGGAGCTGATGCTCCATGCTAAAGATAAAGGTGCTCGT 942
DB 688 TGTACAGCTGTTTACTGCAACATGGAGCTGATGCTCCATGCTAAAGATAAAGGTGCTCGT 747
QY 943 ACCATTACCAATGCTGTTTATGGTCATTATGAAGTAACTGAACCTTTTGGTCAAGCA 1002
DB 748 ACCATTACCAATGCTGTTTATGGTCATTATGAAGTAACTGAACCTTTTGGTCAAGCA 807
QY 1003 TGGTGCCTGTGTAATGCAATGGACTTTGGCAATTCCTCTCTATGAGGAGCTTC 1062
DB 808 TGGTGCCTGTGTAATGCAATGGACTTTGGCAATTCCTCTCTATGAGGAGCTTC 867
QY 1063 TAAGAACAGGGTGAAGTATGTTCTCTCTTAAGTTATGGTGCAGACCCCAACACTGCT 1122

DB 868 TAAGAACAGGGTGAAGTATGTTCTCTTCTTAAAGTTATGGTGCAGACCCCAACACTGCT 927
QY 1123 CAATTGTCAATAAAAGTGTATAGACTTGGCTCCACACACAGATTAAAGAAAGATT 1182
DB 928 CAATTGTCAATAAAAGTGTATAGACTTGGCTCCACACACAGATTAAAGAAAGATT 987
QY 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTCAAGCTGCACGAGAGCTGATGTTACTCG 1242
DB 988 AGCATATGAATTTAAAGGCCACTCGTTGCTCAAGCTGCACGAGAGCTGATGTTACTCG 1047
QY 1243 AATCAAAAAACATCTCTCTGGAATGCTGAATTTCAAGCATCTCAACACATGAAC 1302
DB 1048 AATCAAAAAACATCTCTCTGGAATGCTGAATTTCAAGCATCTCAACACATGAAC 1107
QY 1303 AGCATTTGATTTGCTGCTGCTGCTATCCATATCCCAAGAAAGCAATATATGTAAGTGT 1362
DB 1108 AGCATTTGATTTGCTGCTGCTGCTATCCATATCCCAAGAAAGCAATATATGTAAGTGT 1167
QY 1363 GCTAAGAAAAAGGACCAACATCAATGAAGACATAAAGAAATTTCTTGACTCTCTGCGAGT 1422
DB 1168 GCTAAGAAAAAGGACCAACATCAATGAAGAACTAAAGAAATTTCTTGACTCTCTGCGAGT 1227
QY 1423 GGCATCTGAGAAAGCTCATAATGATTTGTTGAAAGTAGTGTGAAACATGAAGCAAAAGT 1482
DB 1228 GGCATCTGAGAAAGCTCATAATGATTTGTTGAAAGTAGTGTGAAACATGAAGCAAAAGT 1287
QY 1483 TAATGCTCTGATTAATCTTGTGTGAGACTTCTCTACAGAGCTGCATATTTGTTGCTCATCT 1542
DB 1288 TAATGCTCTGATTAATCTTGTGTGAGACTTCTCTACAGAGCTGCATATTTGTTGCTCATCT 1347
QY 1543 ACAACCTGCGCCCTACTCTCTGAGCTATGGGTGTGATCCTTAACATTTATCCCTTCAGGG 1602
DB 1348 ACAACCTGCGCCCTACTCTCTGAGCTATGGGTGTGATCCTTAACATTTATCCCTTCAGGG 1407
QY 1603 CTTTACTGCTTTTACAGATGGGAAATGAAATGTACAGCAACTCTCTCAAGAGGATATCTC 1662
DB 1408 CTTTACTGCTTTTACAGATGGGAAATGAAATGTACAGCAACTCTCTCAAGAGGATATCTC 1467
QY 1663 ATTAGGTAATTCAGAGGACAGACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGA 1722
DB 1468 ATTAGGTAATTCAGAGGACAGACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGA 1527
QY 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCA 1782
DB 1528 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCA 1587
QY 1783 GTCTACACCCTTCAATTTGCGAGTGGGTATACAGAGTGTCCGTGGTGAATATCTGCT 1842
DB 1588 GTCTACACCCTTCAATTTGCGAGTGGGTATACAGAGTGTCCGTGGTGAATATCTGCT 1647
QY 1843 ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCTTTGTACCTTTGCACATGC 1902
DB 1648 ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCTTTGTACCTTTGCACATGC 1707
QY 1903 ATGTTCTTATGGACATTTGAAGTTGCGAAGTCTTTGTTTAAACATGGAGCAGTAA 1962
DB 1708 ATGTTCTTATGGACATTTGAAGTTGCGAAGTCTTTGTTTAAACATGGAGCAGTAA 1767
QY 1963 TGTAGCTGATTTATGGAATTTACACCTTTTACATGAAGAGAGCAAGAAAGAAATATGA 2022
DB 1768 TGTAGCTGATTTATGGAATTTTACACCTTTTACATGAAGAGAGCAAGAAAGAAATATGA 1827
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCACAAAAAAGAGGATGAAA 2082
DB 1828 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCACAAAAAAGAGGATGAAA 1887
QY 2083 TACTCTCTTGGATCTTGTGTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 2142
DB 1888 TACTCTCTTGGATCTTGTGTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 1947
QY 2143 TGCAGCTTTCTAGATGCTGCCAAGAGGCTTTGTTAGCCAGAGTGAAGAGTTGCTTTC 2202
DB 1948 TGCAGCTTTCTAGATGCTGCCAAGAGGCTTTGTTAGCCAGAGTGAAGAGTTGCTTTC 2007

Qy	2203	TCCTGATAATGTAAATTCGCCGATACCCAAAGGAGACATTTCAACACCTTTACATTAGC	2262
Db	2008	TCCTGATAATGTAAATTCGCCGATACCCAAAGGAGACATTTCAACACCTTTACATTAGC	2067
	2263	AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTAA	2322
	2068	AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTAA	2127
Qy	2323	TGCCCAAGACAAGGAGGACTTATTCCCTTTACATAATGCAGCATCTTTACGGGCATCTAGA	2382
Db	2128	TGCCCAAGACAAGGAGGACTTATTCCCTTTACATAATGCAGCATCTTTACGGGCATCTAGA	2187
Qy	2383	TGTAGCAGCTCTACTAATAAAGATATAATGCATGTCTCAATGCCACGACAAATGGGCTTT	2442
Db	2188	TGTAGCAGCTCTACTAATAAAGATATAATGCATGTCTCAATGCCACGACAAATGGGCTTT	2247
Qy	2443	CACACCTTTGCAGGAGACGCCAAAGAGGAGACAGACAGCTTTTGTGCTTTTGTGCTAGC	2502
Db	2248	CACACCTTTGCAGGAGACGCCAAAGAGGAGACAGACAGCTTTTGTGCTTTTGTGCTAGC	2307
Qy	2503	CCATGGAGCTGACCCGACCTTTAAAANTCAGSAAGGACAAACACCTTTACATTTAGTTTC	2562
Db	2308	CCATGGAGCTGACCCGACCTTTAAAANTCAGSAAGGACAAACACCTTTACATTTAGTTTC	2367
Qy	2563	AGCGGATGATGTCAGCGCTTTCTGACAGACGCCATGCCCCATCTGCTCGCCCTCTTG	2622
Db	2368	AGCAGATGATGTCAGCGCTTTCTGACAGACGCCATGCCCCATCTGCTCGCCCTCTTG	2427
Qy	2623	TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTC	2682
Db	2428	TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTC	2487
Qy	2683	TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGCACACATTACTGG	2742
Db	2488	TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGCACACATTACTGG	2547
Qy	2743	GAGTTTTTCAGAACTGCTCTCAGTAGTTAGTTTCAAGTGGAAACAGAGGGTGCTTCCAGTTT	2802
Db	2548	GAGTTTTTCAGAACTGCTCTCAGTAGTTAGTTTCAAGTGGAAACAGAGGGTGCTTCCAGTTT	2607
Qy	2803	GGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG	2862
Db	2608	GGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG	2667
Qy	2863	ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGATGTATTAGTTGA	2922
Db	2668	ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGATGTATTAGTTGA	2727
Qy	2923	GATGGGCACAAAGGAGCTGAAGGAGATTGGAAATCAATGCTTTATGGACATAGGCACAAACT	2982
Db	2728	GATGGGCACAAAGGAGCTGAAGGAGATTGGAAATCAATGCTTTATGGACATAGGCACAAACT	2787
Qy	2983	AATTTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGTCTTTAACCCATATTTAACTTT	3042
Db	2788	AATTTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGTCTTTAACCCATATTTAACTTT	2847
Qy	3043	GAACACCTCTGGTAGTGGAAACAATTTCTATAGATCTGTCTCCTGATGATTAAGAGTTTCA	3102
Db	2848	GAACACCTCTGGTAGTGGAAACAATTTCTATAGATCTGTCTCCTGATGATTAAGAGTTTCA	2907
Qy	3103	GTCTGTGGAGGAAGAGATGCAAAAGTACAGTTCGAGAGCACAGAGATGGAGTCAATGCAGG	3162
Db	2908	GTCTGTGGAGGAAGAGATGCAAAAGTACAGTTCGAGAGCACAGAGATGGAGTCAATGCAGG	2967
Qy	3163	TGGAATCTTCAACAGATACAAATATTTCTCAAGATTCAGAGGTTTGTGAACAAGAACTATG	3222
Db	2968	TGGAATCTTCAACAGATACAAATATTTCTCAAGATTCAGAGGTTTGTGAACAAGAACTATG	3027
Qy	3223	GGAAAGATACACTACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCAATGCCAATGA	3282
Db	3028	GGAAAGATACACTACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCAATGCCAATGA	3087

Qy	3283	ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATATATCCAAAGGCTTTGATGA	3344
Db	3088	ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATATATCCAAAGGCTTTGATGA	3147
Qy	3343	AAGCATGCGTACATAGGTGGTATGTTTCGAGCTGGCATTTATTTGCTGAAAACCTCTTC	3402
Db	3148	AAGCATGCGTACATAGGTGGTATGTTTCGAGCTGGCATTTATTTGCTGAAAACCTCTTC	3207
Qy	3403	CAAAAGCAATCAATATGTATATGAAATGGAGAGAGTACTGGGTGTCAGAGTTCCAAAGA	3462
Db	3208	CAAAAGCAATCAATATGTATATGAAATGGAGAGAGTACTGGGTGTCAGAGTTCCAAAGA	3267
Qy	3463	CAGATCTGTTAGCATTTGCCACAGGCAGCTGCTCTTTTGGCCGGGTAACTTTGGGAAAGTC	3522
Db	3268	CAGATCTGTTAGCATTTGCCACAGGCAGCTGCTCTTTTGGCCGGGTAACTTTGGGAAAGTC	3327
Qy	3523	TTTCTCTGCAGCTTCAGTGCAATGAAATGGCAATCTCCTCCAGGTCACTCACTCAGTCAC	3582
Db	3328	TTTCTCTGCAGCTTCAGTGCAATGAAATGGCAATCTCCTCCAGGTCACTCACTCAGTCAC	3387
Qy	3583	TGGTAGGCCAGTGTAATAGCCCTAGCAATAGCTGAATATGTTATTACAGAGGAGAACA	3642
Db	3388	TGGTAGGCCAGTGTAATAGCCCTAGCAATAGCTGAATATGTTATTACAGAGGAGAACA	3447
Qy	3643	GGCTATTCTCGAGTATTTAAATTAATCTTACCAGATATATGAGGCTGAAGGTATGTCGATGG	3702
Db	3448	GGCTATTCTCGAGTATTTAAATTAATCTTACCAGATATATGAGGCTGAAGGTATGTCGATGG	3507
Qy	3703	A	3703
Db	3508	A	3508
RESULT 12			
AAC66823			
ID	AAC66823 standard; cDNA; 493 BP.		
XX	AAC66823;		
XX	27-FEB-2001 (first entry)		
XX	Human tankyrase II coding sequence SEQ ID NO: 1.		
XX	Human; tankyrase II; telomere length; signal transduction; ss.		
OS	Homo sapiens.		
XX	Location/Qualifiers		
PH	1..4002		
FT	/*tag= a		
FT	/transl_except= (pos:595..597,aa:Xaa)		
FT	/transl_except= (pos:625..630,aa:XaaXaa)		
FT	/transl_except= (pos:637..639,aa:Xaa)		
FT	/transl_except= (pos:694..696,aa:Xaa)		
FT	/transl_except= (pos:724..726,aa:Xaa)		
FT	/transl_except= (pos:730..732,aa:Xaa)		
FT	/transl_except= (pos:763..765,aa:Xaa)		
FT	/transl_except= (pos:790..792,aa:Xaa)		
FT	/transl_except= (pos:2176..2178,aa:Xaa)		
FT	/transl_except= (pos:2380..2382,aa:Xaa)		
FT	/transl_except= (pos:2398..2400,aa:Xaa)		
FT	/transl_except= (pos:2464..2469,aa:XaaXaa)		
FT	/transl_except= (pos:3163..3165,aa:Xaa)		
FT	/note= "Xaa=unknown"		
FT	/product= "tankyrase II"		
FT	/partial		
XX	WO2000061813-A1.		
XX	19-OCT-2000.		
XX	10-APR-2000; 2000WO-US09558.		
XX			

/transl_except= (pos:689..691,aa:Xaa)
 /transl_except= (pos:95..697,aa:Xaa)
 /transl_except= (pos:701..703,aa:Xaa)
 /transl_except= (pos:737..742,aa:XaaXaa)
 /transl_except= (pos:767..772,aa:XaaXaa)
 /transl_except= (pos:2057..2059,aa:Xaa)
 /transl_except= (pos:2231..2233,aa:Xaa)
 /transl_except= (pos:2237..2239,aa:Xaa)
 /transl_except= (pos:2261..2263,aa:Xaa)
 /transl_except= (pos:2279..2281,aa:Xaa)
 /transl_except= (pos:2345..2350,aa:XaaXaa)
 /transl_except= (pos:3035..3037,aa:Xaa)
 /transl_except= (pos:3860..3862,aa:Xaa)
 /transl_except= (pos:3881..3883,aa:Xaa)
 /transl_except= (pos:3890..3892,aa:Xaa)
 /transl_except= (pos:3896..3898,aa:Xaa)
 /transl_except= (pos:3941..3943,aa:Xaa)
 /transl_except= (pos:4001..4003,aa:Xaa)
 /transl_except= (pos:4025..4030,aa:XaaXaa)
 /transl_except= (pos:4118..4120,aa:Xaa)
 /transl_except= (pos:4133..35,aa:Xaa)
 /transl_except= (pos:4172..4174,aa:Xaa)
 /transl_except= (pos:4190..4192,aa:Xaa)
 /transl_except= (pos:4199..4201,aa:Xaa)
 /transl_except= (pos:4238..4240,aa:Xaa)
 /transl_except= (pos:4256..4258,aa:Xaa)
 /transl_except= (pos:4268..4270,aa:Xaa)
 /transl_except= (pos:4283..4285,aa:Xaa)
 /transl_except= (pos:4316..4318,aa:Xaa)
 /note= "Xaa-unknown. The sequence contains a
 number of in-frame termination codons, given as Xaa
 in the protein sequence"
 /product= "tankyrase II"
 /partial

WO200061813-A1.

19-OCT-2000.

10-APR-2000; 2000WO-US09558.

09-APR-1999; 99US-0128577.

13-APR-1999; 99US-0129123.

(GERO-) GERON CORP.

Morin GB, Funk WD, Piatyszek MA;

WPI; 2000-679503/66.

P-PSDB; AAB27210.

Novel mammalian Tankyrase II polypeptide and the polynucleotide
 encoding the polypeptide useful for modulating or maintaining telomere
 length, replicative capacity, apoptosis, chromosome packing or gene
 expression

Disclosure; Fig 3; 52pp; English.

The present sequence is a version of the human tankyrase II coding
 sequence. Its protein is thought to be involved in signal transduction in
 the cell, and to have binding activity for other telomere-associated
 proteins. It is possible that it plays a role in the regulation of
 telomere length, thus affecting the replicative ability of the cell. The
 protein is useful for ribosylating target proteins, for determining
 tankyrase II binding activity in a sample, and for modulating telomere
 length in a cell.

Sequence 4296 BP; 1161 A; 820 C; 925 G; 1033 T; 357 other;

Query Match

Best Local Similarity 78.2%; Score 2968.2; DB 21; Length 4296;

Matches 3020; Conservative 7; Mismatches 381; Indels 12; Gaps 2;

QY 384 CCCTCTGGCTTTATCATTAAGGTTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGA 443
 Db 475 CCCGCTGCATYCCCGCANGTTTGGCGGAAAGACTTANTTAAATATTTGCTTCANA 534
 QY 444 ATGGTGCAAGTCCCAAGCAGCTGATGATGGGGCCCTTATTCCTCTTCATAAATGCATGCT 503
 Db 535 ATGGTGCAAAATGYCAANACCGTCATNATGGGGCCCTTATTCCTCTTCATAAATGCATGCT 594
 QY 504 CTCTTGGTCATGCTGAAGTAGTCAATCTCTTTGGCAGCATGGTGCAGACCCCAATGCTC 563
 Db 595 CTTTGGTCMTGCTRAAANTATCNCTCTTTTGCNACATNGTCANAMCCCAATGCTC 654
 QY 564 GAGATAATTTGGAATATATACCTCTCCATGAAGCTGCAATTAAGAGAAAGATTTGATGTTT 623
 Db 655 GAGATAATTTGGAATATATACCTCTCCNATCAAGCTCAATTAAGAGAAAGATTTGANNNTT 714
 QY 624 GCATTTGCTTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAG 683
 Db 715 GCATTTGCTTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAG 774
 QY 684 CATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGGCTTACTGTGTAATATAAGAAAGATG 743
 Db 775 CATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGGCTTACTGTGTAATATAAGAAAGATG 834
 QY 744 AACTCTTAGAAAGTGCAGAGTGGCAATGAAGAAAGAAATGATGGCTCTACTCACACCAT 803
 Db 835 AACTCTTAGAAAGTGCAGAGTGGCAATGAAGAAAGAAATGATGGCTCTACTCACACCAT 894
 QY 804 TAAATGTCAACTGCCAGCAAGTGAAGGAGCAAGTCAACTCCATTCATTCATTTGGCAGCAG 863
 Db 895 TAAATGTCAACTGCCAGCAAGTGAAGGAGCAAGTCAACTCCATTCATTCATTTGGCAGCAG 954
 QY 864 GATATAACAGAGTAAAGATTGTACAGCTGTACTGCAACATGGAGCTGATGTCATGCTA 923
 Db 955 GATATAACAGAGTAAAGATTGTACAGCTGTACTGCAACATGGAGCTGATGTCATGCTA 1014
 QY 924 AGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCAATTATGAAGTAA 983
 Db 1015 AAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCAATTATGAAGTAA 1074
 QY 984 CTGAACCTTTTGGTCAAGCATGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTC 1043
 Db 1075 CTGAACCTTTTGGTCAAGCATGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTC 1134
 QY 1044 CTCTTCATGAGGCAGCTTCTTAAGAACAGAGTTGAAGTATGTTCTCTCTCTTAAAGTTATG 1103
 Db 1135 CTCTTCATGAGGCAGCTTCTTAAGAACAGAGTTGAAGTATGTTCTCTCTTAAAGTTATG 1194
 QY 1104 GTGCAGACCCACACACTGCTCAATTGTACAAATAAAGTGCATATAGACTTGGCTCCACAC 1163
 Db 1195 GTGCAGACCCACACACTGCTCAATTGTACAAATAAAGTGCATATAGACTTGGCTCCACAC 1254
 QY 1164 CACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTGTTGCTGCAGACTGCAC 1223
 Db 1255 CACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTGTTGCTGCAGACTGCAC 1314
 QY 1224 GAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGC 1283
 Db 1315 GAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGC 1374
 QY 1284 ATCTCTAAACACATGAACACAGCATTCGATGCTGCTGCATCTCCATATCCCAAGAA 1343
 Db 1375 ATCTCTAAACACATGAACACAGCATTCGATGCTGCTGCATCTCCATATCCCAAGAA 1434
 QY 1344 AGCAAAATGTGAACCTGTTGCTTAAGAAAGGAGCAACATCAATGAAGAGACTAAAGAAT 1403
 Db 1435 AGCAAAATGTGAACCTGTTGCTTAAGAAAGGAGCAACATCAATGAAGAGACTAAAGAAT 1494
 QY 1404 TCTTGACTCTCTGCACGTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGG 1463
 Db 1495 TCTTGACTCTCTGCACGTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGG 1554
 QY 1464 TGAACATGAACGAAAGGTTAATGCTCTGGATTAATCTTGGTCAGACTTCTCTACACAGAG 1523

Db	1555	 TGAACATGAAGCAAAAGGTTAATGCTCTGGATAATCTTTGGTCAGACATCTCTACACAGAG	1614
Qy	1524	CTGCATATTCTGGTCAATCTACAAACCTGCGGCTACTCCTCAGCTATTGGGTGTGATCCCTA	1583
Db	1615	CTGCATATTCTGGTCAATCTACAAACCTGCGGCTACTCCTCAGCTATTGGGTGTGATCCCTA	1674
Qy	1584	ACATTATATCCCTTCAGGGCTTTACTTGTCTTTACAGATGGGAAATGAAAATGTACAGCAAC	1643
Db	1675	ACATTATATCCCTTCAGGGCTTTACTTGTCTTTACAGATGGGAAATGAAAATGTACAGCAAC	1734
Qy	1644	TCCTCCAAGAGGTATCTCATTTAGTGAATTCAGAGGCACAGACAAATTTGCTGGAAGCTG	1703
Db	1735	TCCTCCAAGAGGTATCTCATTTAGTGAATTCAGAGGCACAGACAAATTTGCTGGAAGCTG	1794
Qy	1704	CAAAAGCTGCAGATGTCGAAACTGTAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGCA	1763
Db	1795	CAAAAGCTGCAGATGTCGAAACTGTAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGCA	1854
Qy	1764	GAGACATTTGAAGGGCTCAGTCTACACCACCTTCATTTTTCAGCTGGGTATAAACAGTGT	1823
Db	1855	GAGACATTTGAAGGGCTCAGTCTACACCACCTTCATTTTTCAGCTGGGTATAAACAGTGT	1914
Qy	1824	CCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTCATCTCTAAAGATAAAGGAGGCC	1883
Db	1915	CCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTCATCTCTAAAGATAAAGGAGGCC	1974
Qy	1884	TTGTACCTTTTGCACAATGCATGCTCTTATGAGACATTATGAAGTTGCAGAACTCTCTGTTA	1943
Db	1975	TTGTACCTTTTGCACAATGCATGCTCTTATGAGACATTATGAAGTTGCAGAACTCTCTGTTA	2034
Qy	1944	AACATGGAGCAGTAGTTAATGTAGCTGATTTATGAAATTTACACCTTTACATGAAGCAG	2003
Db	2035	AACATGGAGCAGTAGTTAATGTAGCTGATTTATGAAATTTACACCTTTACATGAAGCAG	2094
Qy	2004	CAGCAAAAGGAAATATGAAATTTGCAAACTTCTCCTCCAGCATGGTGCAGACCTTACCA	2063
Db	2095	CAGCAAAAGGAAATATGAAATTTGCAAACTTCTCCTCCAGCATGGTGCAGACCTTACCA	2154
Qy	2064	AAAAAAACAGGGATGAAATACTCCTTTTGGATCTTGTATAAGATGGAGATACAGATATTC	2123
Db	2155	AAAAAAACAGGGATGAAATACTCCTTTTGGATCTTGTATAAGATGGAGATACAGATATTC	2214
Qy	2124	AAGATCTGCTTAGGGAGATGCAGCTTTTGTAGATGCTGCCAAAGAGGGTTGTTTAGCCA	2183
Db	2215	AAGATCTGCTTAGGGAGATGCAGCTTTTGTAGATGCTGCCAAAGAGGGTTGTTTAGCCA	2274
Qy	2184	GAGTGAAGAAGTTGCTCTTCTCTGATAATGTAAATTTGCCGCGATACCCAGGCGAGCATT	2243
Db	2275	GAGTGAAGAAGTTGCTCTTCTCTGATAATGTAAATTTGCCGCGATACCCAGGCGAGCATT	2334
Qy	2244	CAACACCTTTTACATTTAGCAGCTGTATAATTAATTTAGAGTTGCAGAGTATTGTTTAC	2303
Db	2335	CAACACCTTTTACATTTAGCA--GGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	2391
Qy	2304	AACACGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATCGAG	2363
Db	2392	NN	2451
Qy	2364	CATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAGTATAATGTCATGTCGAATG	2423
Db	2452	NN	2511
Qy	2424	CCACGGCAAAATGGCTTTCACACCTTTGCACGAAGACGCCCAAGGACACACAGC	2483
Db	2512	NN	2571
Qy	2484	TTTGTGCTTTGTCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGACAAA	2543
Db	2572	NN	2631
Qy	2544	CACCTTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGCACAGCAGCCATGCC	2603

QY 3675 TTATGAGGCTGAAGGTATGTCGATGATAAATAGTATTATTTAAGAACTAATTCACCT 3734
|||||
Db 3772 TTATGAGGCTGAAGGTATGTCGATGATAAATAGTATTATTTAAGAACTAATTCACCT 3831
|||||
QY 3735 GAACCTAAATCA'CAAGCAGCAGTGGCTCTACGTTTACTCCTTTGCTGAAAAAAA 3794
|||||
Db 3832 GAACCTAAATCA'CAAGCAGCAGTGGCTCTACGTTTACTCCTTTGCTGAAAAAAA 3891
|||||
RESULT 14
AAC89706
ID AAC89706 standard; cdna; 5005 BP.
XX AC AAC89706;
XX 08-MAR-2001 (first entry)
XX Human adult T-cell leukaemia cDNA MO-BC-203.
DE XX
XX Human; cytostatic; gene therapy; vaccine; breast cancer;
KW T cell leukaemia; cancer associated antigen; ss.
XX
XX Homo sapiens.
OS
XX WO200072021-A2.
PN 30-NOV-2000.
PD 24-MAY-2000; 2000WO-US14391.
PF 26-MAY-1999; 99US-0320092.
PR (LUDW-) LUDWIG INST CANCER RES.
XX
XX Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;
PI Sahin U, Piruendschuh M, Old LJ, Knuth A, Jager E;
XX
XX WPI; 2001-032067/04.
DR P-PSDB; AAB48574.
XX
XX Diagnosing disorder characterized by expression of human cancer
PT associated antigen precursor, involves detecting interaction of an
PT agent with novel NA Group 1 nucleic acid molecule encoding the antigen
PT precursor -
XX
XX Example 8; Page 82-84; 87pp; English.
PS
XX The present sequence is given in a specification relating to nucleic
CC acids and encoded polypeptides which are cancer associated antigens
CC expressed in patients afflicted with breast cancer and/or T cell
CC leukaemia. A disorder characterised by expression of human cancer
CC associated antigen may be diagnosed by contacting a biological sample
CC with an agent that specifically binds to the nucleic acid that encodes
CC the antigen, complexed with human leukocyte antigen molecule, and
CC determining the interaction between the agent and the nucleic acid.
CC Cancer associated antigens, the nucleotides encoding them, antibodies
CC against them and the pharmaceutical compositions comprising them are
CC useful for diagnosing, monitoring and treating the diseases characterised
CC by the expression of one or more cancer associated antigens.
XX
SQ Sequence 5005 BP; 1492 A; 917 C; 1015 G; 1571 T; 10 other;

Query Match 73.1%; Score 2774.6; DB 22; Length 5005;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2810; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY 978 AAGTAACGAACCTTTGGTCAAGCATGGTGGCTGTGTAAATGCAATGGACTTGTGGCAAT 1037
|||||
Db 1 AAGTAACGAACCTTTGGTCAAGCATGGTGGCTGTGTAAATGCAATGGACTTGTGGCAAT 60
|||||
QY 1038 TCACTCCTCTTCATGAGGAGCTTCTAAGACAGGGTTGAAGTATGTTCTCTCTCTTAA 1097
|||||

Db 61 TCACTCCTCTTCATGAGGAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAA 120
|||||
QY 1098 GTTATGTCGAGACCCCAACACTGCTCAATTTGTCACAATAAAAGTGCATATAGCTTGGCTC 1157
|||||
Db 121 GTTATGTCGAGACCCCAACACTGCTCAATTTGTCACAATAAAAGTGCATATAGCTTGGCTC 180
|||||
QY 1158 CCACACCACAGTAAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1217
|||||
Db 181 CCACACCACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 240
|||||
QY 1218 CTGCACGAGAGCTGATGTTACTCGAATCAAAAAACATCTCTCTGGAATGTTGAAT 1277
|||||
Db 241 CTGCACGAGAGCTGATGTTACTCGAATCAAAAAACATCTCTCTGGAATGTTGAAT 300
|||||
QY 1278 TCAGCATCTCAACACATGAACACAGATTGCAATTTGTCGTCGATCTCCATATCCCA 1337
|||||
Db 301 TCAAGCATCTCAACACATGAACACAGATTGCAATTTGTCGTCGATCTCCATATCCCA 360
|||||
QY 1338 AAGAAAAGCAATATGTAAGTGTGCTAAGAAAAGAGCAACATCAATGAAAAGACTA 1397
|||||
Db 361 AAGAAAAGCAATATGTAAGTGTGCTAAGAAAAGAGCAACATCAATGAAAAGACTA 420
|||||
QY 1398 AAGAAATCTTTGACTCCTCTGCACGTGGCATCTGAGAAGCTCATAATGATGTTTGAAG 1457
|||||
Db 421 AAGAAATCTTTGACTCCTCTGCACGTGGCATCTGAGAAGCTCATAATGATGTTTGAAG 480
|||||
QY 1458 TAGTGTGAAACATGAAGCAAAAGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1517
|||||
Db 481 TAGTGTGAAACATGAAGCAAAAGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 540
|||||
QY 1518 ACAGAGCTGCATATTTGGTGCATCTACAAACCTGCCGCTACTCCTGAGCTATGGGTGTG 1577
|||||
Db 541 ACAGAGCTGCATATTTGGTGCATCTACAAACCTGCCGCTACTCCTGAGCTATGGGTGTG 600
|||||
QY 1578 ATCTTAACATATATCCTTTTACAGGGCTTTACTGCTTTACAGATGGGAAATGAAATGTAC 1637
|||||
Db 601 AUCCTAACATATATCCTTTTACAGGGCTTTACTGCTTTACAGATGGGAAATGAAATGTAC 660
|||||
QY 1638 AGCAACTCCTCCCAAGAGGGTATCTCATTAGTAAATTCAGAGGCAGACAGCAATGGCTGG 1697
|||||
Db 661 AGCAACTCCTCCCAAGAGGGTATCTCATTAGTAAATTCAGAGGCAGACAGCAATGGCTGG 720
|||||
QY 1698 AAGTCGAAAGGCTGGAGATGTCGAAACTGTGAAAAAAGTGTGACCTGTTTACAGAGTGTCA 1757
|||||
Db 721 AAGTCGAAAGGCTGGAGATGTCGAAACTGTGAAAAAAGTGTGACCTGTTTACAGAGTGTCA 780
|||||
QY 1758 ACTGCAGAGACATTTGAAGGGCGTCAGTCTACACACATTCATTTTGCAGCTGGGTATAA 1817
|||||
Db 781 ACTGCAGAGACATTTGAAGGGCGTCAGTCTACACACATTCATTTTGCAGCTGGGTATAA 840
|||||
QY 1818 GAGTGTCCGTTGGTGAATATCTGCTACAGCATGAGCTGATGTCATGCTTAAAGATAAAG 1877
|||||
Db 841 GAGTGTCCGTTGGTGAATATCTGCTACAGCATGAGCTGATGTCATGCTTAAAGATAAAG 900
|||||
QY 1878 GAGGCTTTGACCTTTGCAACAATGCAATGTTCTTATGACATTTATGAAGTTGCAAACTTC 1937
|||||
Db 901 GAGGCTTTGACCTTTGCAACAATGCAATGTTCTTATGACATTTATGAATTTGCAAACTTC 960
|||||
QY 1938 TTGTTAAACATGAGCAGTATGTTAAATGATGCTGATTTATGAAATTTTACACCTTTTACATG 1997
|||||
Db 961 TTGTTAAACATGAGCAGTATGTTAAATGATGCTGATTTATGAAATTTTACACCTTTTACATG 1020
|||||
QY 1998 AAGCAGCAGCAAAAGGAAATATGAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2057
|||||
Db 1021 AAGCAGCAGCAAAAGGAAATATGAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 1080
|||||
QY 2058 CTACCAAAAAAAGAGGATGGAATTAATCTCCTTTGGATCTTGTAAAGATGGAGATACAG 2117
|||||
Db 1081 CTACCAAAAAAAGAGGATGGAATTAATCTCCTTTGGATCTTGTAAAGATGGAGATACAG 1140
|||||
QY 2118 ATATTCAAGATCTGTTAGGGGAGATGACCTTTTGTGATGCTGCTGCAAGAGGGTTGTT 2177
|||||

Db	1141	ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT	1200
Qy	2178	TAGCCAGAGTGAAGAAGTTGCTTCTCTGATAATGTAAATTCGCGGATACCAAGCA	2237
Db	1201	TAGCCAGAGTGAAGAAGTTGCTTCTCTGATAATGTAAATTCGCGGATACCAAGCA	1260
Qy	2238	GACATTCACACCTTTTACATTTAGCAGCTGGTTATATAATTTAGAAGTTGCAGATATT	2297
Db	1261	GACATTCACACCTTTTACATTTAGCAGCTGGTTATATAATTTAGAAGTTGCAGATATT	1320
Qy	2298	TGTTTACAACAGCGAGCTGATGTGAATGCCCAAGACAAGGAGGACTTATTCCTTTACATA	2357
Db	1321	TGTTTACAACAGCGAGCTGATGTGAATGCCCAAGACAAGGAGGACTTATTCCTTTACATA	1380
Qy	2358	ATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGATATAATGCAATGTG	2417
Db	1381	ATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGATATAATGCAATGTG	1440
Qy	2418	TCATGCCACGGACAATGGCTTTTCACACCTTTGACAGAGCAGCCCAAGAGGACGAA	2477
Db	1441	TCATGCCACGGACAATGGCTTTTCACACCTTTGACAGAGCAGCCCAAGAGGACGAA	1500
Qy	2478	CACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAG	2537
Db	1501	CACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAG	1560
Qy	2538	GACAAACACCTTTTAGATTTTAGTTTTCAGCGGATGATGTCAGCGCTCTTCTGCACAGCAGCCA	2597
Db	1561	GACAAACACCTTTTAGATTTTAGTTTTCAGCAGATGATGTCAGCGCTCTTCTGCACAGCAGCCA	1620
Qy	2598	TGCCCCCATCTGCTCTGCCCCCTCTGTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCC	2657
Db	1621	TGCCCCCATCTGCTCTGCCCCCTCTGTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCC	1680
Qy	2658	CAGAGGCCATGCAGATGCTCTCTCTTCAGTGTCATCTAGCCCATCAAGGCTTTCTGTCAG	2717
Db	1681	CAGAGGCCATGCAGATGCTCTCTCTTCAGTGTCATCTAGCCCATCAAGGCTTTCTGTCAG	1739
Qy	2718	CCAGCAGCTTTGACAACTTATCTGGGAGTTTTTCAGAACTCTCTTCAGTAGTTAGTTCAA	2777
Db	1740	CCAGCAGCTTTGACAACTTATCTGGGAGTTTTTCAGAACTCTCTTCAGTAGTTAGTTCAA	1797
Qy	2778	GTGGAACAGAGGGTGCTCCAGTTTGGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCA	2837
Db	1798	GTGGAACAGAGGGTGCTCCAGTTTGGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCA	1857
Qy	2838	TAACTCAATTCGTAAGGAATCTTTGGACTTTGAGCACCCTAATGGATATATTTGAGAGAGAAC	2897
Db	1858	TAACTCAATTCGTAAGGAATCTTTGGACTTTGAGCACCCTAATGGATATATTTGAGAGAGAAC	1917
Qy	2898	AGATCACTTTTGGATGATTAAGTTGAGATGGGCACAAGGAGCTGAAGGAGATTGGAATCA	2957
Db	1918	AGATCACTTTTGGATGATTAAGTTGAGATGGGCACAAGGAGCTGAAGGAGATTGGAATCA	1977
Qy	2958	ATGCTTATGGACATAGGCACAACTAATTAAGGAGTGCAGAGACTTATCTCCGGACAAC	3017
Db	1978	ATGCTTATGGACATAGGCACAACTAATTAAGGAGTGCAGAGACTTATCTCCGGACAAC	2037
Qy	3018	AAGTCTTAAACCATATTTAACTTTTGAACACCTCTGGTAGTGGACAATTTCTTATAGATC	3077
Db	2038	AAGTCTTAAACCATATTTAACTTTTGAACACCTCTGGTAGTGGACAATTTCTTATAGATC	2097
Qy	3078	TGCTCTCTGATGAAGAAGTTTCACTCTGTGGAGGAGAGATGCAAGATACAGTTCCGAG	3137
Db	2098	TGCTCTCTGATGAAGAAGTTTCACTCTGTGGAGGAGAGATGCAAGATACAGTTCCGAG	2157
Qy	3138	AGCACAGAGATGGAGGTCATGACAGTGGAACTCTTCAACAGATACAAATATCTCTCAAGATT	3197
Db	2158	AGCACAGAGATGGAGGTCATGACAGTGGAACTCTTCAACAGATACAAATATCTCTCAAGATT	2217
Qy	3198	AGAAGGTTTGTAAACAAGAACTATGGAAAGATACACTCACCGGAGAAAGAACTTTCTG	3257
Db	2218	AGAAGGTTTGTAAACAAGAACTATGGAAAGATACACTCACCGGAGAAAGAACTTTCTG	2277

Qy	3258	AAGAAAACCAACCAATGCCAATGAACGAATGCTATTTCACTAGGGTCTCCTTTCTGGAATG	33117
Db	2278	AAGAAAACCAACCAATGCCAATGAACGAATGCTATTTCACTAGGGTCTCCTTTCTGGAATG	2337
Qy	3318	CNATTATTCCAAGAAGGCTTTGATGAAGAAGCATGCGTACATAGGTGGTATGTTTGGAGCTG	33777
Db	2338	CAATTATCCAAAGGCTTTGATGAAGAAGCATGCGTACATAGGTGGTATGTTTGGAGCTG	23997
Qy	3378	GCATTTATTTTGGCTGAAAACCTCTCCAAAGCAATCAATATGTTATGGAATTTGGAGGAG	34377
Db	2398	GCATTTATTTTGGCTGAAAACCTCTCCAAAGCAATCAATATGTTATGGAATTTGGAGGAG	24577
Qy	3438	GTACTGGGCTGCCAGTTTCAAAAGACAGATCTTGTTCACATTTGCCACAGCGACGTGCTCT	34977
Db	2458	GTACTGGGCTGCCAGTTTCAAAAGACAGATCTTGTTCACATTTGCCACAGCGACGTGCTCT	25177
Qy	3498	TTTCCGGGTAAOCTTTGGGAAAGCTTTTCCTGCAGTTCAGTCAATGAAATATGGCACATT	35577
Db	2518	TTTCCGGGTAAOCTTTGGGAAAGCTTTTCCTGCAGTTCAGTCAATGAAATATGGCACATT	25777
.Qy	3558	CTCCTCCAGGTCAATCACTCAGTTCACCTGCTAGCCACAGTGAATGGCCTAGCATTAGCTG	36177
Db	2578	CTCCTCCAGGTCAATCACTCAGTTCACCTGCTAGCCACAGTGAATGGCCTAGCATTAGCTG	26377
Qy	3618	AATATGTTATTTACAGAGGAGAACAGGCTTATCTCGAGTATTTTAATTAATTAACACAGATTA	36777
Db	2638	AATATGTTATTTACAGAGGAGAACAGGCTTATCTCGAGTATTTTAATTAATTAACACAGATTA	26977
Qy	3678	TGAGGCTTGAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACCTAAATTCACACGAA	37377
Db	2698	TGAGGCTTGAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACCTAAATTCACACGAA	27577
Qy	3738	CCTAAATCAATCAAGACAGAGTGGCCTCTAGCTTTTACTCCTTTGCTGAAAAAAA	3794
Db	2758	CCTAAATCAATCAAGACAGAGTGGCCTCTAGCTTTTACTCCTTTGCTGAAAAAAA	2814

RESULT 15	
AAAF63919	
ID	AAAF63919 standard; DNA; 2971 BP.
XX	
AC	AAF63919;
XX	
DT	05-APR-2001 (first entry)
XX	
DE	Human tankyrase2 clone consensus SEQ ID NO: 92.
XX	
KW	Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW	inflammatory disorder; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200100849-A1.
XX	
PD	04-JAN-2001.
XX	
PF	28-JUN-2000; 200WO-US17827.
XX	
PR	29-JUN-1999; 99US-0141582.
XX	
PA	(ICOS-) ICOS CORP.
XX	
PI	Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX	
DR	WPI; 2001-102896/11.
XX	
PT	New tankyrase2 polypeptides, useful for treating conditions media-
PT	poly(adenosine diphosphate-ribose) polymerase activity e.g. cancer
PT	inflammatory and autoimmune disorders -
XX	
PS	Example 1; Page 152-153; 242pp; English.
XX	

CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
SQ

Sequence 2971 BP; 915 A; 555 C; 642 G; 859 T; 0 other;

Query Match 64.3%; Score 2439.8; DB 22; Length 2971;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1352 TGTCAACTGTTGTGTAAGAGGAGCAACATCAATGAAAGACTAAAGATTCCTTGACT 1411
DB 1 TGTCAACTGTTGTGTAAGAGGAGCAACATCAATGAAAGACTAAAGATTCCTTGACT 60
QY 1412 CCTCGACCTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAAGTAGTGGTGAACAT 1471
DB 61 CCTCGACCTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAAGTAGTGGTGAACAT 120
QY 1472 GAAGCAAGGTTAATGCTCTGGATAATCTTGCTCAGACTTCTCTACAGAGCTGCATAT 1531
DB 121 GAAGCAAGGTTAATGCTCTGGATAATCTTGCTCAGACTTCTCTACAGAGCTGCATAT 180
QY 1532 TGTGGTGCATACAAACCTGCCGCCCTACTCCTGAGCTATGGGTGTCATCCTTAACATTATA 1591
DB 181 TGTGGTGCATACAAACCTGCCGCCCTACTCCTGAGCTATGGGTGTCATCCTTAACATTATA 240
QY 1592 TCCCTTCAGGGCTTTACTGTTTACAGATGGGAAATGAAATGTACAGCAACTCCCTCCAA 1651
DB 241 TCCCTTCAGGGCTTTACTGTTTACAGATGGGAAATGAAATGTACAGCAACTCCCTCCAA 300
QY 1652 GAGGTATCTCATTAGTAAATTCAGAGGAGACAGACAATTCCTGAAAGCTGCAAGGCT 1711
DB 301 GAGGTATCTCATTAGTAAATTCAGAGGAGACAGACAATTCCTGAAAGCTGCAAGGCT 360
QY 1712 GGAGATGTCGAAACTGTAAATAACTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATT 1771
DB 361 GGAGATGTCGAAACTGTGTAAATAACTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATT 420
QY 1772 GAAGGGCGTCAGTCTACACCCTCACTTTTTCAGCTGGGTATTAACAGAGTGCCTGGTG 1831
DB 421 GAAGGGCGTCAGTCTACACCCTCACTTTTTCAGCTGGGTATTAACAGAGTGCCTGGTG 480
QY 1832 GAATATCTGCTACAGCATGGAGCTGATGTCATGCTTAAAGATAAAGAGGCCCTTGACCT 1891
DB 481 GAATATCTGCTACAGCATGGAGCTGATGTCATGCTTAAAGATAAAGAGGCCCTTGACCT 540
QY 1892 TTGCACAATGCATGTTCTTATGGACATTTAAGTTGCAGAACTTCTTGTAAACATGGA 1951
DB 541 TTGCACAATGCATGTTCTTATGGACATTTAAGTTGCAGAACTTCTTGTAAACATGGA 600
QY 1952 GCAGTAGTAAATGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAA 2011
DB 601 GCAGTAGTAAATGCTGATTTATGGAATTTACACCTTTTACATGAAGCAGCAGCAAAA 660
QY 2012 GGAATAATGAAATTTGCAAACTTCTGCTCCAGCATGTTGTCAGACCTACCAAAAAAAC 2071
DB 661 GGAATAATGAAATTTGCAAACTTCTGCTCCAGCATGTTGTCAGACCTACCAAAAAAAC 720
QY 2072 AGGGATGGAATACCTCTTGGATCTGTTTAAAGATGGAGATACAGATATTAAGATCTG 2131
DB 721 AGGGATGGAATACCTCTTGGATCTGTTTAAAGATGGAGATACAGATATTAAGATCTG 780
QY 2132 CTTAGGGAGATGACGCTTTGCTAGATCTGCTCAAGAGGGTGTGTTACCGAGAGTGAAG 2191
DB 781 CTTAGGGAGATGACGCTTTGCTAGATCTGCTCAAGAGGGTGTGTTTACCGAGAGTGAAG 840
QY 2192 AAGTTGCTCTCTCCTGATAATGTAATGCCCGCATACCCAGGCGAGACATTCAACACCT 2251
DB 1921 CATGCCAATGAACGATGCTATTTCATGGGTCTCTCTTTGTGAATGCAATTATCCACAAA 1980

DB 841 AAGTTGCTCTCTCTCTGATAAATGTAAATTCGCCGATACCCAGGCGAGACATTCAACACCT 900
QY 2252 TTACATTTAGCAGCTGGTTATAAATTTAGAAAGTTGCAGAGTATTGTTTAAACACCGGA 2311
DB 901 TTACATTTAGCAGCTGGTTATAAATTTAGAAAGTTGCAGAGTATTGTTTAAACACCGGA 960
QY 2312 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAAATGCAGCATCTTAC 2371
DB 961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAAATGCAGCATCTTAC 1020
QY 2372 GGGCATGTAGATGTAGCAGCTCTACTATAAAGTATAATATGATGTGTCATGCCACGGAC 2431
DB 1021 GGGCATGTAGATGTAGCAGCTCTACTATAAAGTATAATGATGTGTCATGCCACGGAC 1080
QY 2432 AAATGGGCTTTCACACCTTTGACAAAGAGCCCAAAAGGAGCAACACAGCTTTTGTGCT 2491
DB 1081 AAATGGGCTTTCACACCTTTGACAAAGAGCCCAAAAGGAGCAACACAGCTTTTGTGCT 1140
QY 2492 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAATAATCAGGAGGACAAACACCTTTA 2551
DB 1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAATAATCAGGAGGACAAACACCTTTA 1200
QY 2552 GATTTAGTTTTCAGCGGATGATGTACAGCGCTCTCTGACAGAGCCATGCCCCCATCTGCT 2611
DB 1201 GATTTAGTTTTCAGCGGATGATGTACAGCGCTCTCTGACAGAGCCATGCCCCCATCTGCT 1260
QY 2612 CTGCCCTCTTGTACAAGCCCTCAAGTGTCTCAATGGTGTGAGAAAGCCAGGAGCCACTGCA 2671
DB 1261 CTGCCCTCTTGTACAAGCCCTCAAGTGTCTCAATGGTGTGAGAAAGCCAGGAGCCACTGCA 1320
QY 2672 GATGCTCTCTCTCAGGTCCTAGCCCATCAAGCCCTTCTGACAGCCAGCAGCTTTTGAC 2731
DB 1321 GATGCTCTCTCTCAGGTCCTAGCCCATCAAGCCCTTCTGACAGCCAGCAGCTTTTGAC 1380
QY 2732 AACTTATCTGGGAGTTTTCAGAACTGTCCTCAGTAGTATTAGTCAAGTGGAAACAGAGGT 2791
DB 1381 AACTTATCTGGGAGTTTTCAGAACTGTCCTCAGTAGTATTAGTCAAGTGGAAACAGAGGT 1440
QY 2792 GCTTCCAGTTTGGAGAAAGAGGTTTCCAGAGTATGATTTTACATAACTCAATTCGTA 2851
DB 1441 GCTTCCAGTTTGGAGAAAGAGGTTTCCAGAGTATGATTTTACATAACTCAATTCGTA 1500
QY 2852 AGGAATCTTGGACTTGGACACCTAATGGATATATTGAGAGAGAACAGATCACTTTGGAT 2911
DB 1501 AGGAATCTTGGACTTGGACACCTAATGGATATATTGAGAGAGAACAGATCACTTTGGAT 1560
QY 2912 GTATTAGTTGAGATGGGCAACAGGAGCTGAAGAGATTTGGAATCAATGCTTTATGGACAT 2971
DB 1561 GTATTAGTTGAGATGGGCAACAGGAGCTGAAGAGATTTGGAATCAATGCTTTATGGACAT 1620
QY 2972 AGGCACAACTTAATTAAGGAGTCGAGAGCTTATCTCCGACAAACAGGCTTTAAACCA 3031
DB 1621 AGGCACAACTTAATTAAGGAGTCGAGAGCTTATCTCCGACAAACAGGCTTTAAACCA 1680
QY 3032 TATTTAACTTTGAACACCTCTGTTAGTGGAAACAAATTTTATAGATCTGCTCTCTGATGAT 3091
DB 1681 TATTTAACTTTGAACACCTCTGTTAGTGGAAACAAATTTTATAGATCTGCTCTCTGATGAT 1740
QY 3092 AAAGAGTTTCAGTCTGTTGGAGGAGAGATGCAAGATACAGTTTCGAGAGCACAGATGGA 3151
DB 1741 AAAGAGTTTCAGTCTGTTGGAGGAGAGATGCAAGATACAGTTTCGAGAGCACAGATGGA 1800
QY 3152 GGTGATGAGTGGAACTTCTCAACAGATCAATATTTCTCAAGATTGAGAAGGTTTGTAAAC 3211
DB 1801 GGTGATGAGTGGAACTTCTCAACAGATCAATATTTCTCAAGATTGAGAAGGTTTGTAAAC 1860
QY 3212 AAGAACTATGGGAAAGATACACTCAGCAGGAGAAAGAGTTTCTGAAGAAACCAAC 3271
DB 1861 AAGAACTATGGGAAAGATACACTCAGCAGGAGAAAGAGTTTCTGAAGAAACCAAC 1920
QY 3272 CATGCCAATGAACGATGCTATTTCATGGGTCTCTCTTTGTGAATGCAATTATCCACAAA 3331
DB 1921 CATGCCAATGAACGATGCTATTTCATGGGTCTCTCTTTGTGAATGCAATTATCCACAAA 1980

Qy 3332 GGCTTTGATGAAAGGCATCGGTACATAGTGGTGTATGTTGGAGCTGGCAATTTATTTGCT 3391
Db 1981 GGCTTTGATGAAAGGCATCGGTACATAGTGGTGTATGTTGGAGCTGGCAATTTATTTGCT 2040
Qy 3392 GAAACTCTTCCAAAGCAATCAATATGATATGGAATTGGAGGAGGTACTGGGTGTCCA 3451
Db 2041 GAAACTCTTCCAAAGCAATCAATATGATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100
Qy 3452 GTTCACAAAGACAGATCTTGTACATTTGCCACAGGAGGTGCTCTTTTGGCCGGGTAACC 3511
Db 2101 GTTCACAAAGACAGATCTTGTACATTTGCCACAGGAGGTGCTCTTTTGGCCGGGTAACC 2160
Qy 3512 TTGGGAAAGTCTTTCCCTGCAGTTCAGTGCATGAAATGGCACATTTCTCCTCCAGGTCAAT 3571
Db 2161 TTGGGAAAGTCTTTCCCTGCAGTTCAGTGCATGAAATGGCACATTTCTCCTCCAGGTCAAT 2220
Qy 3572 CACTCAGTCACTGTAGGCCAGGTAAATGGCTAGCATTTAGTGAATATGTTATTAC 3631
Db 2221 CACTCAGTCACTGTAGGCCAGGTAAATGGCTAGCATTTAGTGAATATGTTATTAC 2280
Qy 3632 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTTACCAGATTTATGAGGCCTGAAGGT 3691
Db 2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTTACCAGATTTATGAGGCCTGAAGGT 2340
Qy 3692 ATGTCGATGATAAATAGTATTTTAAAGAACTAATTCCTGAACTGAACTAAATCATCAA 3751
Db 2341 ATGTCGATGATAAATAGTATTTTAAAGAACTAATTCCTGAACTGAACTAAATCATCAA 2400
Qy 3752 AGCAGCAGTGGCCTCTAGCTTTTACTCTTTTGTGAAAAAAA 3794
Db 2401 AGCAGCAGTGGCCTCTAGCTTTTACTCTTTTGTGAAAAAAA 2443

Search completed: February 11, 2003, 13:10:06
Job time : 752.206 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 12:17:58 ; Search time 115,212 Seconds
(without alignments)
10107.057 Million cell updates/sec

Title: US-09-843-159B-1
Perfect score: 3797
Sequence: 1 cttgaagacactggatttc.....ccttctgctgaaaaa 3797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3382.8	89.1	4512	4	US-09-350-982C-3
2	3291.8	86.7	3498	4	US-09-350-982C-4
3	1813.4	47.8	4134	4	US-09-196-387-1
4	1449.6	38.2	4491	4	US-09-196-387-7
5	1297.8	34.2	4657	4	US-09-196-387-9
6	121.2	3.2	523	4	US-09-196-387-12
7	93	2.4	5235	1	US-09-031-485-35
8	93	2.4	5235	1	US-09-031-485-36
9	93	2.4	5235	1	US-08-847-429A-35
10	93	2.4	5235	1	US-08-847-429A-36
11	93	2.4	5235	3	US-09-065-474-35
12	93	2.4	5235	3	US-09-065-474-36
13	93	2.4	5235	4	US-09-557-034-35
14	93	2.4	5235	4	US-09-557-034-36
15	93	2.4	5503	1	US-09-031-485-32
16	93	2.4	5503	1	US-09-031-485-33
17	93	2.4	5503	1	US-08-847-429A-32
18	93	2.4	5503	1	US-08-847-429A-34
19	93	2.4	5503	3	US-09-065-474-32
20	93	2.4	5503	3	US-09-065-474-33
21	93	2.4	5503	4	US-09-557-034-32
22	93	2.4	5503	4	US-09-557-034-34
23	83.4	2.2	1056	3	US-09-065-474-138
24	83.4	2.2	1056	4	US-09-065-474-140
25	83.4	2.2	1056	4	US-09-557-034-138
26	83.4	2.2	1056	4	US-09-557-034-140
27	78.8	2.1	3454	4	US-09-082-059-1

28	74.2	2.0	2505	4	US-09-291-839-3	Sequence 3, Appli
29	74.2	2.0	3025	4	US-09-291-839-1	Sequence 1, Appli
30	73.8	1.9	909	1	US-09-031-485-25	Sequence 25, Appl
31	73.8	1.9	909	1	US-09-031-485-26	Sequence 26, Appl
32	73.8	1.9	909	1	US-08-847-429A-25	Sequence 25, Appl
33	73.8	1.9	909	1	US-08-847-429A-26	Sequence 26, Appl
34	73.8	1.9	909	3	US-09-065-474-25	Sequence 25, Appl
35	73.8	1.9	909	3	US-09-065-474-26	Sequence 26, Appl
36	73.8	1.9	909	4	US-09-557-034-25	Sequence 25, Appl
37	73.8	1.9	909	4	US-09-557-034-26	Sequence 26, Appl
38	73.8	1.9	911	1	US-09-031-485-22	Sequence 22, Appl
39	73.8	1.9	911	1	US-09-031-485-24	Sequence 24, Appl
40	73.8	1.9	911	1	US-08-847-429A-22	Sequence 22, Appl
41	73.8	1.9	911	1	US-08-847-429A-24	Sequence 24, Appl
42	73.8	1.9	911	3	US-09-065-474-22	Sequence 22, Appl
43	73.8	1.9	911	3	US-09-065-474-24	Sequence 24, Appl
44	73.8	1.9	911	4	US-09-557-034-22	Sequence 22, Appl
45	73.8	1.9	911	4	US-09-557-034-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-09-350-982C-3
; Sequence 3, Application US/09350982C
; Patent No. 6455290
; GENERAL INFORMATION:
; APPLICANT: Berthelsen, Jens
; APPLICANT: Toma, Salvatore
; APPLICANT: Isacchi, Antonella
; TITLE OF INVENTION: Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods R
; TITLE OF INVENTION: Same
; FILE REFERENCE: PHRM-0043
; CURRENT APPLICATION NUMBER: US/09/350,982C
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 4512
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: PCR Primers
; NAME/KEY: misc_feature
; LOCATION: (1124)..(1124)
; OTHER INFORMATION: n is any nucleic acid
; NAME/KEY: misc_feature
; LOCATION: (2672)..(2672)
; OTHER INFORMATION: n is any nucleic acid
US-09-350-982C-3

Query Match	89.1%	Score 3382.8;	DB 4;	Length 4512;
Best Local Similarity	99.7%	Pred. No. 0;		
Matches 3381;	Conservative	6;	Mismatches	5;
			Indels	0;
			Gaps	0;
QY	403	AGGTTTGGCGGAAGACACCTAGTTGAATATTTGCTTCAGATGGTGCAAGTGTCCAAGC	462	
Db	220	AGGTTTGGCGGAAGACACCTAGTTGAATATTTGCTTCAGATGGTGCAAGTGTCCAAGC	279	
QY	463	ACGTGATGATGGGGGCTTATTCCTTCATAATGATGATGCTCTTTGGTCATGCTCAAGT	522	
Db	280	ACGTGATGATGGGGGCTTATTCCTTCATAATGATGATGCTCTTTGGTCATGCTCAAGT	339	
QY	523	AGTCAATCTCTTTTGGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC	582	
Db	340	AGTCAATCTCTTTTGGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC	399	
QY	583	TCCTCTCCATGAAGCTGCAATTAAGAAAGATTTGATGTTTGCATGCTGTGTACAGCA	642	
Db	400	TCCTCTCCATGAAGCTGCAATTAAGAAAGATTTGATGTTTGCATGCTGTGTACAGCA	459	
QY	643	TGGAGCTGAGCCCAACCATCGAATACAGATGAAGACAGCATTGGATTAGCAGATCC	702	

|||||
Db 460 TGGAGCTGAGCAACCATCCGAAATACAGATGAAGGACAGCATTTGGATTTAGCAGATCC 519
QY 703 ATCTGCCAAGCAGTCTTACTGCTGGAATATAAGAAAGATGAACCTTTAGAAAGTGCCAG 762
Db 520 ATCTGCCAAGCAGTCTTACTGCTGGAATATAAGAAAGATGAACCTTTAGAAAGTGCCAG 579
QY 763 GAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTTAAATGTCAACTGCCACGC 822
Db 580 GAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTTAAATGTCAACTGCCACGC 639
QY 823 AAGTATGCCAAGTCAACTCCATTAATGATGGCAGCAGGATATAACAGAGTAAAGAT 882
Db 640 AAGTATGCCAAGTCAACTCCATTAATGATGGCAGCAGGATATAACAGAGTAAAGAT 699
QY 883 TGTACAGCTGTTTACTGCAACATGGAGCTCATGCTCATGCTAAAGATAAAGGTGATCTGCT 942
Db 700 TGTACAGCTGTTTACTGCAACATGGAGCTGATGCTCAATGAAGATAAAGGTGATCTGCT 759
QY 943 ACCATTACAAATGGCTGCTTATGGTCATTAATGAAGTAACCTGAACCTTTTGGTCAAGCA 1002
Db 760 ACCATTACAAATGGCTGCTTATGGTCATTAATGAAGTAACCTGAACCTTTTGGTCAAGCA 819
QY 1003 TGGTGCCTGTAAATGCAATGGACTTGGCAATTCACCTCTCTTCATGAGGCAGCTTC 1062
Db 820 TGGTGCCTGTAAATGCAATGGACTTGGCAATTCACCTCTCTTCATGAGGCAGCTTC 879
QY 1063 TAAGAACAGGGTGAAGTATGTTCTCTTAAAGTTATGCTGAGAGCCCAACACTGCT 1122
Db 880 TAAGAACAGGGTGAAGTATGTTCTCTTAAAGTTATGCTGAGAGCCCAACACTGCT 939
QY 1123 CAATTGTCAATAAAGTGCTATAGACTTGGCTCCACACACAGTAAAAAGAAAGATT 1182
Db 940 CAATTGTCAATAAAGTGCTATAGACTTGGCTCCACACACAGTAAAAAGAAAGATT 999
QY 1183 AGCATATGAATTTAAGGCCACTGCTGCTGCAAGCTGCACGAGAGCTGATGTTACTCG 1242
Db 1000 AGCATATGAATTTAAGGCCACTGCTGCTGCAAGCTGCACGAGAGCTGATGTTACTCG 1059
QY 1243 AATCAAAAACATCTCTCTGGAATGCTGAATTTCAAGCATCTCAAAACACATGAAC 1302
Db 1060 AATCAAAAACATCTCTCTGGAATGCTGAATTTCAAGCATCTCAAAACACATGAAC 1119
QY 1303 AGCATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1362
Db 1120 AGCATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1179
QY 1363 GCTAAGAAAGGAGCAACATCAATCAAGACTAAAGATTTTGAATTTTGAATTTTGAATTT 1422
Db 1180 GCTAAGAAAGGAGCAACATCAATCAAGACTAAAGATTTTGAATTTTGAATTTTGAATTT 1239
QY 1423 GGCATCTGAGAAGCTCATPAATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGT 1482
Db 1240 GGCATCTGAGAAGCTCATPAATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGT 1299
QY 1483 TAATGCTCTGGATAATCTTGGTCTGAGACTTCTCTACAGAGCTGCTATGTTGGTCACT 1542
Db 1300 TAATGCTCTGGATAATCTTGGTCTGAGACTTCTCTACAGAGCTGCTATGTTGGTCACT 1359
QY 1543 ACAAACTCCGCTACTCTCAGCTATGGTGTGATCTCTACATTTATATCCCTTCAGGG 1602
Db 1360 ACAAACTCCGCTACTCTCAGCTATGGTGTGATCTCTACATTTATATCCCTTCAGGG 1419
QY 1603 CTTTACTGTTTACAGATGGGAAATGAAATGTACAGCAACTCCCTCAAGAGGGTATCTC 1662
Db 1420 CTTTACTGTTTACAGATGGGAAATGAAATGTACAGCAACTCCCTCAAGAGGGTATCTC 1479
QY 1663 ATTAGGTAATTCAGAGGCAGACAGCAATGCTGGAAGCTGCAAGGCTGGAGATGTCGA 1722
Db 1480 ATTAGGTAATTCAGAGGCAGACAGCAATGCTGGAAGCTGCAAGGCTGGAGATGTCGA 1539
QY 1723 AACTGTAAAAAATCTGTACTGTTTCAGAGTGTCAACTGCACAGACATTTGAAGGGGTCA 1782
|||||

Db 1540 AACTGTAAAAAATCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCTCA 1599
QY 1783 GTCTACACACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCT 1842
Db 1600 GTCTACACACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCT 1659
QY 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCCTTGTAACCTTTGCACAATGC 1902
Db 1660 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGRRGCCCTTGTAACCTTTGCACAATGC 1719
QY 1903 ATGTTCTTATGACATTTAAGATTGCAACAATCTTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1720 ATGTTCTTATGACATTTAAGATTGCAACAATCTTTGTTAAACATGGAGCAGTAGTTAA 1779
QY 1963 TGTAGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAAGGAAATATGA 2022
Db 1780 TGTAGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAAGGAAATATGA 1839
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGCCCTACCAAAAAAAGGAGGTGAAA 2082
Db 1840 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGCCCTACCAAAAAAAGGAGGTGAAA 1899
QY 2083 TACTCCTTTGGATCTGTTTAAAGATGGATACAGATATTCAGATCTCTTAGGGAGA 2142
Db 1900 TACTCCTTTGGATCTGTTTAAAGATGGATACAGATATTCAGATCTCTTAGGGAGA 1959
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAAGTTGCTTC 2202
Db 1960 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAAGTTGCTTC 2019
QY 2203 TCCGTGATAATGTAATTTGCCGGATACCCAAAGCAGACATTCACAACTTTACATTTAGC 2262
Db 2020 TCCGTGATAATGTAATTTGCCGGATACCCAAAGCAGACATTCACAACTTTACATTTAGC 2079
QY 2263 AGCTGTTTATATAATTTAGAAAGTTCAGAGTATTTGTTTACACAGGAGCTGATGTGA 2322
Db 2080 AGCTGTTTATATAATTTAGAAAGTTCAGAGTATTTGTTTACACAGGAGCTGATGTGA 2139
QY 2323 TGCCCAAGACAAAGAGGAGCTTATTCCTTTACATATATGAGCATCTTACGGGCATGTAGA 2382
Db 2140 TGCCCAAGACAAAGAGGAGCTTATTCCTTTACATATATGAGCATCTTACGGGCATGTAGA 2199
QY 2383 TGTAGAGCTCTACTAATAAAGTATTAATGCATGTGTCAATGTCACGCAAAATGGGCTTT 2442
Db 2200 TGTAGAGCTCTACTAATAAAGTATTAATGCATGTGTCAATGTCACGCAAAATGGGCTTT 2259
QY 2443 CACACCTTTGACGAGCAGCAGCCCAAAAGGAGCAGACACAGCTTTGTTGCTGTAGC 2502
Db 2260 CACACCTTTGACGAGCAGCAGCCCAAAAGGAGCAGACACAGCTTTGTTGCTGTAGC 2319
QY 2503 CCATGGAGCTGACCCGACCTTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2320 CCATGGAGCTGACCCGACCTTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2379
QY 2563 AGCGATGATGTACGGCTCTCTGACAGCAGCATGCCCATGCTGCTCTGCTGCTGCTGCTG 2622
Db 2380 AGCGATGATGTACGGCTCTCTGACAGCAGCATGCCCATGCTGCTCTGCTGCTGCTGCTG 2439
QY 2623 TTTACAAAGCCTCAAGTGTCTCAATGGTGTGAGAGCCAGGAGCAGTGCAGATGCTCTC 2682
Db 2440 TTTACAAAGCCTCAAGTGTCTCAATGGTGTGAGAGCCAGGAGCAGTGCAGATGCTCTC 2499
QY 2683 TTTAGGTCCATCTAGCCCATCAAGCCTTTCTGACGAGCAGCTCTTTGACAACCTTTATCTGG 2742
Db 2500 TTTAGGTCCATCTAGCCCATCAAGCCTTTCTGACGAGCAGCTCTTTGACAACCTTTATCTGG 2559
QY 2743 GAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGGTCTTCCAGTTT 2802
Db 2560 GAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGGTCTTCCAGTTT 2619
QY 2803 GGAGAAAAAGAGGGTTCAGAGGAGTAGTTTTAGCATAACTCAATTCCTAAGGATCTTGG 2862
Db 2620 GGAGAAAAAGAGGGTTCAGAGGAGTAGTTTTAGCATAACTCAATTCCTAAGGATCTTGG 2679
|||||

Qy	1063	TAAGAACAGGGTTGAAGTATGTTCTCTCTTTAAGTTAAGTGCAGACCCCAACACTGCT	1122
Db	858	TAAGAACAGGGTTGAAGTATGTTCTCTCTTTAAGTTAAGTGCAGACCCCAACACTGCT	917
Qy	1123	CAATTGTCACAATAAAAGTGCCTATACACTTGGCTCCCAACACACAGTTTAAAGAAGATT	1182
Db	918	CAATTGTCACAATAAAAGTGCCTATACACTTGGCTCCCAACACACAGTTTAAAGAAGATT	977
Qy	1183	AGCATATGAATTTAAAGGCCACTCGTTCGTGCAAGCTGCACGAGAAGCTGATGTTACTCG	1242
Db	978	AGCATATGAATTTAAAGGCCACTCGTTCGTGCAAGCTGCACGAGAAGCTGATGTTACTCG	1037
Qy	1243	AATCAAAAACATCTCTCTCTGGAATGOTGAATTTCAAGCAATCCTCAACACATGAAC	1302
Db	1038	AATCAAAAACATCTCTCTCTGGAATGOTGAATTTCAAGCAATCCTCAACACATGAAC	1097
Qy	1303	AGCATTCGATTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAAATATGTGAACCTGTT	1362
Db	1098	AGCATTCGATTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAAATATGTGAACCTGTT	1157
Qy	1363	GCTAAGAAAAGGAGCAACATCAATGAAAAGACTTAAAGAAATTTTGACTCCTCTGCACGT	1422
Db	1158	GCTAAGAAAAGGAGCAACATCAATGAAAAGACTTAAAGAAATTTTGACTCCTCTGCACGT	1217
Qy	1423	GGCATCTGCAGAAAGCTCATTAATGATGTTTGAAGTAGTGGTGAACAATGAAGCAAAAGT	1482
Db	1218	GGCATCTGCAGAAAGCTCATTAATGATGTTTGAAGTAGTGGTGAACAATGAAGCAAAAGT	1277
Qy	1483	TAATGCTCTGGATAAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATCTGTGCTCATCT	1542
Db	1278	TAATGCTCTGGATAAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATCTGTGCTCATCT	1337
Qy	1543	ACAAACCTGCGGCTACTCTGAGCTATGGGTGATGCTTAAACATTAATCCCTTCACGG	1602
Db	1338	ACAAACCTGCGGCTACTCTGAGCTATGGGTGATGCTTAAACATTAATCCCTTCACGG	1397
Qy	1603	CTTTTACGTCTTTACAGATGGGAATGAAATGTACAGCAACTCCTCCAGAGGATCTCTC	1662
Db	1398	CTTTTACGTCTTTACAGATGGGAATGAAATGTACAGCAACTCCTCCAGAGGATCTCTC	1457
Qy	1663	ATTAGGTAAATTCAGAGGCAGACAGACAATTCGTGCAAGCTGCAAGGCTGGAGATGTCGA	1722
Db	1458	ATTAGGTAAATTCAGAGGCAGACAGACAATTCGTGCAAGCTGCAAGGCTGGAGATGTCGA	1517
Qy	1723	AACGTGAAAAAACTGTGTACTGTTACAGAGTGTCAACGTGCAGAGACATPGAAGGGCGTCA	1782
Db	1518	AACGTGAAAAAACTGTGTACTGTTACAGAGTGTCAACGTGCAGAGACATPGAAGGGCGTCA	1577
Qy	1783	GTCTACACCACTTCATTTTGCAGCTGGGTATACAGAGAGTGCCTGGTGAATATCTGCT	1842
Db	1578	GTCTACACCACTTCATTTTGCAGCTGGGTATACAGAGAGTGCCTGGTGAATATCTGCT	1637
Qy	1843	ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGAGCCCTGTACTTTTGACAAATGC	1902
Db	1638	ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGAGCCCTGTACTTTTGACAAATGC	1697
Qy	1903	ATGTTCTTATGGACATATGAAGTTGCAGACTCTCTGTTAAACATGGACGAGTAGTTAA	1962
Db	1698	ATGTTCTTATGGACATATGAAGTTGCAGACTCTCTGTTAAACATGGACGAGTAGTTAA	1757
Qy	1963	TGTAGCTGATTATGAAATTTACACCTTTACATGAAGCAGCAACCAAGAAATATGA	2022
Db	1758	TGTAGCTGATTATGAAATTTACACCTTTACATGAAGCAGCAACCAAGAAATATGA	1817
Qy	2023	AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACAAAAAACAACAGGATGAAA	2082
Db	1818	AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACAAAAAACAACAGGATGAAA	1877
Qy	2083	TACTCCTTTGGATCTGTGTTAAAGATGGAGATACAGATTAATCAAGATCTGCTTAGGGAGA	2142
Db	1878	TACTCCTTTGGATCTGTGTTAAAGATGGAGATACAGATTAATCAAGATCTGCTTAGGGAGA	1937

QY	2143	TG	CAGCTTTGCTAGATGCTGCCAAGAAGGGTGTGTTATGCCAGAGTGAAGAAGTGTGCTTC	2202
Db	1938	TG	CAGCTTTGCTAGATGCTGCCAAGAAGAGTGTGTTAGCCAGAGTGAAGAAGTGTCTTC	1997
QY	2203	TC	CTGATAATGTAATTCGCGCATACCCAAGCAGACATTCACACCTTTACATTTAGC	2262
Db	1998	TC	CTGATAATGTAATTCGCGCATACCCAAGCAGACATTCACACCTTTACATTTAGC	2057
QY	2263	AG	TGTTTAAATTTAGAAGTTCGAGAGTATTTGTTTAAACACGGAGCTGATGTGAA	2322
Db	2058	AG	TGTTTAAATTTAGAAGTTCGAGAGTATTTGTTTAAACACGGAGCTGATGTGAA	2117
QY	2323	TG	CCCAAGCAAGGAGGACTTATTCCTTTACATAATGCAGACTTTACGGGCAATGAGA	2382
Db	2118	TG	CCCAAGCAAGGAGGACTTATTCCTTTACATAATGCAGACTTTACGGGCAATGAGA	2177
QY	2383	TG	TAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT	2442
Db	2178	TG	TAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT	2237
QY	2443	CAC	ACCTTTGCAGAGCAGGCCAAAAGGACGAACACAGCTTTTGCTTTGTGTGCTPAGC	2502
Db	2238	CAC	ACCTTTGCAGAGCAGGCCAAAAGGACGAACACAGCTTTTGCTTTGTGTGCTPAGC	2297
QY	2503	CC	ATGAGCTGACCCGACCTCTTAAAAATCAGGAAGCACAACACCTTTAGATTTACTTTTC	2562
Db	2298	CC	ATGAGCTGACCCGACCTCTTAAAAATCAGGAAGCACAACACCTTTAGATTTACTTTTC	2357
QY	2563	AG	CGGATGATGCAGCGCTCTCTTGACAGACGCCATGCCGCCATCTGCTCGCCCTCTTG	2622
Db	2358	AG	CGGATGATGTCAGCGCTCTCTTGACAGACGCCATGCCGCCATCTGCTCGCCCTCTTG	2417
QY	2623	TT	CAGCCTCAAGTCTCAATGGTGTGAGAGCCGAGGACCACTGCAGATGCTCTCTC	2682
Db	2418	TT	CAGCCTCAAGTCTCAATGGTGTGAGAGCCGAGGACCACTGCAGATGCTCTCTC	2477
QY	2683	TT	CAGTCCATCTAGCCCATCAAGCCCTTCTGCAGCCAGCAGTCTTGACAACCTTATCTGG	2742
Db	2478	TT	CAGTCCATCTAGCCCATCAAGCCCTTCTGCAGCCAGCAGTCTTGACAACCTTATCTGG	2537
QY	2743	GAG	TTTTTCAGACACTCTCTCAGTAGTTAGTTCAAGTGAACAGAGGGTCTTCCAGTTT	2802
Db	2538	GAG	TTTTTCAGACACTCTCTCAGTAGTTAGTTCAAGTGAACAGAGGGTCTTCCAGTTT	2597
QY	2803	GG	AAAAAGAGGTTCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG	2862
Db	2598	GG	AAAAAGAGGTTCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG	2657
QY	2863	ACT	TGAGCACCTTAATGGATATATTTCAGAGAGAACAGATCACTTTGGATGTATTAGTTGA	2922
Db	2658	ACT	TGAGCACCTTAATGGATATATTTCAGAGAGAACAGATCACTTTGGATGTATTAGTTGA	2717
QY	2923	GAT	GGGCAAGAGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT	2982
Db	2718	GAT	GGGCAAGAGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT	2777
QY	2983	AAT	TAAAGAGTCGAGAGACTTATCTCCGGACAACAAGTCTTAACCCATATTTAACTTT	3042
Db	2778	AAT	TAAAGAGTCGAGAGACTTATCTCCGGACAACAAGTCTTAACCCATATTTAACTTT	2837
QY	3043	GA	ACCTCTGTTAGGAACAATTTTATAGATCTGCTCTGATGATTAAGAGTTTCA	3102
Db	2838	GA	ACCTCTGTTAGGAACAATTTTATAGATCTGCTCTGATGATTAAGAGTTTCA	2897
QY	3103	GT	CTGGAGGAGATGCAAAGTACAGTTTCGAGAGCACAGAGATGGAGGTCATGCAGG	3162
Db	2898	GT	CTGGAGGAGAGATGCAAAGTACAGTTTCGAGAGCACAGAGATGGAGGTCATGCAGG	2957
QY	3163	TG	GAATCTTCAACAGATCAATATTTCTCAAGATTCAAGAGTTTGTAAACAAGAACTATG	3222
Db	2958	TG	GAATCTTCAACAGATCAATATTTCTCAAGATTCAAGAGTTTGTAAACAAGAACTATG	3017
QY	3223	GG	AAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAAACCAACCACTGCCAATGA	3282

3018	Db		GGAAAGATACACTACCGAGAAAAAGATTCTCTGAAGAAACACAAACATGCCAATGA	3077
3283	QY		ACCAATGCCTATTTTCATGGGTCCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA	3342
3078	Db		ACGAATGCCTATTTTCATGGGTCTCCCTTTTGTGAATGCAAAATATCCACAAAGGCTTTGATGA	3137
3343	QY		AAGGCATCGGTACATAGGTGGTGTATTTTGGAGCTGGCATTTATTTTGGCTGAAACCTCTTC	3402
3138	Db		AAGGCATCGGTACATAGGTGGTGTATTTTGGAGCTGGCATTTATTTTGGCTGAAACCTCTTC	3197
3403	QY		CAAAAGCAATCAATATGTATATGGAAATTTGGAGGAGGTACTGGGTGCCAGTTTACAAAGA	3462
3198	Db		CAAAAGCAATCAATATGTATATGGAAATTTGGAGGAGGTACTGGGTGCCAGTTTACAAAGA	3257
3463	QY		CAGATCTTGTACATTTGGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTC	3522
3258	Db		CAGATCTTGTACATTTGGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTC	3317
3523	QY		TTTCCCTGCAGTTCACTGCAATGAAATGGCACATCTCTCCCTCCAGGTCAATCAGTCAGTCAC	3582
3318	Db		TTTCCCTGCAGTTCACTGCAATGAAATGGCACATCTCTCCAGGTCAATCAGTCAGTCAC	3377
3583	QY		TGCTAGGCCCCAGTGTAAATGGCCTAGCATTAAGCTGAATATCTTTATTTACAGAGGAGAACA	3642
3378	Db		TGCTAGGCCCCAGTGTAAATGGCCTAGCATTAAGCTGAATATCTTTATTTACAGAGGAGAACA	3437
3643	QY		GGCTTATCTCTGAGTATTTAAATTACTTACACAGATTATGAGGCTCGAAGGTATGGTCCATGG	3702
3438	Db		GGCTTATCTCTGAGTATTTAAATTACTTACACAGATTATGAGGCTCGAAGGTATGGTCCATGG	3497
3703	QY		A 3703	
3498	Db		A 3498	

РЕЗУЛТАТ 3

US-09-196-387-1
 ; Sequence 1, Application US/09196387
 ; Patent No. 6277613
 ; GENERAL INFORMATION:
 ; APPLICANT: de Lange, Titia
 ; APPLICANT: Smith, Susan
 ; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
 ; TITLE OF INVENTION: OF USE THEREOF
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klauber & Jackson
 ; STREET: 411 Hackensack Avenue, 4th Floor
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/196,387
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/095,225
 ; FILING DATE: June 10, 1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684
 ; TELEX: 133521

QY 1283 CATCCTCAAAACACATGAACAGCATTTGCATTGCTGCTGCATCTCCATATCCAAAGA 1342
D 1557 CAACCGAGTCTCATGAACAGACACTGCATGCTGTGGCTCTCTGCACTCCAAACGT 1616
QY 1343 AAGCAAAATATGTGAATGCTGCTTAAGAAAGAGGACAAACATCAATGAAAGACTAAAGAA 1402
D 1617 AAACAAGTGACAGAAATTTTACTTAGAAAGAGGACAAATGTTAATGAAAAAAATAAGAT 1676
QY 1403 TTCTTCACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAAATGTTGTTGAAGTAGTG 1462
D 1677 TCCATGACCTCCCTGCTGATGTCAGCGCGAAAGAGCCCATAAATGATGTCATGGAAGTTCTG 1736
QY 1463 GTGAACATCAAGCAAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522
D 1737 CATAGCATGGCCCAAGATGAATGACCTGGACACCCCTTGGTCAGACTGCTTTGCAATAGA 1796
QY 1523 GCTGCATATTTGGTGCATCTACAAACCTGCGCCTACTCCTGAGCTATGGTGTGATCCT 1582
D 1797 GCGGCCCTAGCAGGCCACCTGCAGACCTGCGCCTCTGCTGAGTTACGGCTCTGACCCC 1856
QY 1583 AACATATATCCTTCCAGGCTTACTGCTTACAGATGGGAATCAAAATGTACAGCAA 1642
D 1857 TCCATCATCTCTTACAAGGCTTACAGCAGCACAGATGGCAATGAAGCAGTGCAGCAG 1916
QY 1643 CTCCTCCAAGAGGGTATCTCAATAGTAGTAATTCAGAGGCAGACAGCAATTTGCTGGAAGCT 1702
D 1917 ATTCTGAGTGAGAGTACCTATACGTACTCTGATGTTGATGATATCGACTCTTAGAGGCA 1976
QY 1703 GCMAAGCTGGAGATGCGAACTGTAAACCTGTGACTGTGCTGAGAGTGTCAACTGC 1762
D 1977 TCTAAAGCTGGAGACTGTGAAACTGTGAAGCAACTTTCAGAGCTCTCAAAATGTAATGT 2036
QY 1763 AGAGCATTTGAAGGGCTGAGTCAACACCTTCAATTTGCAGCTGGGTATACACAGTG 1822
D 2037 AGAGCTTAGAGGGCGGCTTCCAGCCCTTACACTTCGACAGCGCTACAAACCGCTG 2096
QY 1823 TCCGTGGTGAATATCTGTACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCG 1882
D 2097 TCTGTGTAGAGTACCTGTACACCAGCTGCCGATGCTCCATGCCAAAGAGGCTGC 2156
QY 1883 CTGTGACCTTTCCAAATCATGCTTCTATGACATATGAGATGTCAGAACTCTTGT 1942
D 2157 TTGGTGGCCCTTCATAAGCCCTGTCATGAGCACTATGAGTGGCTGAGCTTTTAGTA 2216
QY 1943 AAACATGGAGCAGTAGTTAATAGCTGATTTATGGAATTTACACCTTTACATGAAGCA 2002
D 2217 AGGCATGGGCTTCTGCAATGTGGCGACTTATGGAATTTACCCCTCTCCATGAAGCA 2276
QY 2003 GCAGCAAAAGGAAATATGAATTTGCAACTTCTGCTCCAGCATGGTGCAGACCCCTACC 2062
D 2277 GCAGCTAAGGAAAGTATGAATCTGCAAGCTCTCTTTTAAACATGGAGCAGATCCAAT 2336
QY 2063 AAAAAAAGAGGATGGAATACTCTTTGGATCTTGTAAAGTGGAGATACAGATAT 2122
D 2337 AAAAAGACAGAGATGGAATACACTTTGGATTTGGTAAAGGAAAGAGACACAGATAT 2396
QY 2123 CAAGATCTGCTTAGGGGAGATGACGCTTGTCTGATGCTGCCAAGAGGTTGTTTAGCC 2182
D 2397 CAGGACTTACTGAAGGGAGTCTGCTTTGTTGGATGCTGCCAAGAGGCTGCTGCCA 2456
QY 2183 AGAGTGAAGAGTCTCTCTGCTGATATGTTAAATGTCGCCGATGACCAAGGACAGAT 2242
D 2457 AGAGTGAAGAGTCTGATACCCAGAGATATCAACTGCAGAGACACCCAGGCGAGAAAT 2516
QY 2243 TCAACACCTTTACATTTAGCAGCTGTTATATATTTAGAGTTGACAGATATTTGTTA 2302
D 2517 TCAACCCCTCTGCACCTGGCAGCGGCTATATATACCTTGAAGTAGCTCAATATCTCTA 2576
QY 2303 CAACAGGAGCTGATGTGAATGCCCAAGCAAGAGGAGCTTATTCCTTTACATAATGCA 2362
D 2577 GAGCATGGAGCTGATGTTAATGGCCAGCAGAGGGTGTAAATTCCTCTTCATATCGG 2636

QY 2363 GCATCTTACGGGCATGTAGATCTAGCAGCTCTACTATATAAGTATATATGATGTGTCAAT 2422
D 2637 GCATCTTATGGGCATGTGTACATAGCGCTTTATTTGATAAAATACAAACAGTGTGTAAAT 2696
QY 2423 GCACAGCAAAATTTGGCTTTTACACACTTTGCACGAAGCAGCCCAAAAGGAGCAACACAG 2482
D 2697 GCAACAGATAAGTGGGGTTTACTCCCTCCATGAAGCAGCCCAAGAAAGAGGAGCCAG 2756
QY 2483 CTTTGTGCTTTGTTGTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGCAAA 2542
D 2757 CTGTGCGCCTCTCTCTAGCGCATGTGACACCCACCCACCATGAAGAACCCAGGAAGGCCAG 2816
QY 2543 ACACCTTTAGATTTAGTTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCC 2602
D 2817 ACCTCTCTGATCTGCAACAGCTGACGATATCAGAGCTTTGCTGATAGATGCTCATGCC 2876
QY 2603 CCAATCTGCTTGCCTCTTTTACAAGCCCTCAAGTGTCTAATGTTGTGAGAAGCCAGGA 2662
D 2877 CCAGAGGCTTACCTACCTCTTTAAACCTCAGG-----CTACT 2915
QY 2663 GCCACTGCAGATGCTCTCTCTCAGGTGCATCTAGCCCATCAAGCCCTTTCTGCAGCCAG 2722
D 2916 GTAGTGAGTGCCTCTCTGATCTCACCAGCATCCACCCCTCTCTGCTCGGCTGCCAG 2975
QY 2723 AGTCTTTGACAACTTATCTGGAGTTTTTTCAGAACTCTCTTCACTAGTTAGTTCAAGTGA 2782
D 2976 AGCATAGACAACTCTACTGGCCCTTTAGCAGAGTTGGCGGTAGGAGGAGCCCTCCAATGCA 3035
QY 2783 ACAGAGGCTGCTTCCAGTTTGGAGAAAAAG-----GAGTTCCAGGAGTAGATTTTAGC 2836
D 3036 GGGATGGCGCGCGGGAACAGAAAGAGGAGGAGAGTGTCTGCTTGATGATCAAT 3095
QY 2837 ATAACATCAATTCGTAAGGAATCTTGACCTTGACACCTAATGATATATTTGAGAGAGAA 2896
D 3096 ATCAGCCAAATTTCTAAAAAGCCTTGGCTTGAACACTTTCGGGATATCTTTGAACAGAA 3155
QY 2897 CAGATCACCTTTGGATGATTTAGTTGAGATGGGACACAGAGCTGAAGAGATTTGGAATC 2956
D 3156 CAGATTACACTAGATGTTGGCTGATATGGTCTGTAAGAGTTGAAAGAAATAGGCATC 3215
QY 2957 AATGCTTATGGACATAGGCACAAACTTAATAAGGAGTGCAGAGACTTATCTCCGACAA 3016
D 3216 AATGATATAGGACCGCCCAAAATTAATCAAAAGAGTAGAAGACTCTTAGTGGACAA 3275
QY 3017 CAAGTCTTAACCCATATTTAACTTTGAACCTCTGGTAGTGGACAAATTTCTTATAGAT 3076
D 3276 CAAGGACCAATCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGTGGAT 3335
QY 3077 CTGCTCTCTGATGATAAGAGTTTTCAGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTCA 3136
D 3336 CTGCTCCAGAGATATAAGAAATATCAGTCAGTGGAGAGAGATGCAAAAGTACTATTCA 3395
QY 3137 GAGCAGAGATGGAGGCTCATGAGTGGAAATCTTCAACAGATACAAATTTCTCAAGATT 3196
D 3396 GAACACAGAGATGGTGAATGTAATGCTGGCGCATCTTCAACAGATACAAATGCTATCGAAT 3455
QY 3197 CAGAGGTTTGAACAAAGAACTATGGGAAAGATACACTCACCAGGAGAAAGAAAGTTCT 3256
D 3456 CAAAAGTTGTCAACAGAAAGTTGAGGAGCGGTTCTGCCACCCAGAGAAAGTGTCT 3515
QY 3257 GAAGAAAACCAACCATGCCCATAAGCAATGCTATTTTCATGGGTCTCTCTTTTGTGAAT 3316
D 3516 GAGGAGATCAACACCATCACAATGAGCGATGTTGTTTCATGGTCTCTCTTTCATTAAT 3575
QY 3317 GCAATTTCCAAAGGCTTTGATGAAAGGCAATGCGTACATAGGTGGTATGTTGGAGCT 3376
D 3576 GCAATTTATCAAGGCTTTGATGAGGACATGATACATAGGAGGAATGTTTGGGGCC 3635
QY 3377 GGCATTTATTTGCTGAAAACCTCTTCCAAAAGCAATCAATATGATATGGAATTCGAGCA 3436
D 3636 GGGATTTATTTGCTGAAAACCTCTTCCAAAAGCAACCAATATGTTTATGGAATTCGAGCA 3695
QY 3437 GGTACTGGGTGTCCAGTTTCACAAAAGACAGATCTTGTATACATTTGCCACAGGAGCTGCTC 3496

INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 5235 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-031-485-36

Query Match 2.4%; Score 93; DB 1; Length 5235;
Best Local Similarity 45.0%; Pred. No. 1.8e-16;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

QY 1373 GGAGCAAACTCAATGAAAGAACTAAAGAAATTCCTGCTCTCTGCGCTGCGCATCTCGAG 1432
DB 5058 GGCAGGATATCAACACATGCAATCGGAATGGCCCTTAATGCAATCTGCGCTCCAAA 4999
QY 1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGTGAAACATGAAGCAAAAGTTAAATGCTCTG 1492
DB 4998 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAGAAAAGCAGATGTTGATGCTGCC 4939
QY 1493 GATAATCTTGGTCAGACITCTCTACAGAGCTGCATATTTGGHCACTTACAAACCTGC 1552
DB 4938 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCAITTTGGCAGGACAAAGAACTAATCGTC 4879
QY 1553 CGCCTACTCTGAGCTATGCGTGTGATCTCTAATCATATATATCCCTTCAGGGCTTTACTGCT 1612
DB 4878 ACAGTACTTGTGAAATGGTCTTAATGTTAACGTACAACTACTAAACGGTTTACACCA 4819
QY 1613 TTACAGATGGGAATGAAATGTACAGCAACTCCCTCAAGAGGGTATCTCATTTAGGTAAT 1672
DB 4818 CTTTACATGGCTGCACAGAAAATCAGGAATCTGTTGACGCTATCTTCTTGCCTCCACAAT 4759
QY 1673 TCAGAGGACAGACAAATTTGCTGAGAGCTGCAAGAGCTGGAGATGT-----CGAAAC 1725
DB 4758 GCAATCAAGCTTTAAGTACAGAAAGCGGTTTACGCCACTGGCAGTTGCCCTTGAACAA 4699
QY 1726 TGTAAAAAACTGTGTACGTTTCAGAGTGTCACTGCAGAGACATTTGAAGGGCGTCAGTC 1785
DB 4698 GGTGACGATCGTGTGCTGCTGTTGCTTGAATATGACACCGCGGAAAGTG-CGCTT 4640
QY 1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGGAATATCTCTACA 1845
DB 4639 GCCAGCACTGCATATGCTGCTAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4580
QY 1846 GCATGAGCTGATGTCATGCTTAAGATAAAGAGGCGCTTGTACCTTTGCCACAATGCATG 1905
DB 4579 AAATGAGCATAACTCGGATGTGACTTCGAAAGCGGCTTTACTCCGCTTCATATCCCGC 4520
QY 1906 TTCTTTATGACATTTAAGTTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTAAATGT 1965
DB 4519 TCATATGAATGAGACGTAGCAGCACTGCTACTCGAAAGGAGGCAATGTGAATTA 4460
QY 1966 AGCTGATTTATGAAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAT 2025
DB 4459 CCAAGCGAGACATAACATAAGTCGGTTACACGTTGCAACAAAATGGGTCGTACAAACAT 4400
QY 2026 TTGCAAACTTCTGCTCCAGATGTTGCGACAGCCCTTCCAAAAAACAAGGATGGAATAC 2085
DB 4399 GGTTTCGTTATTTGGCTCATGCGGCGCGCTTAATGACTGTCACACACGTGATTTACTAAC 4340
QY 2086 TCCTTTGGATCTGTTTAAAGATGGAGATACAGAT---ATTCAAGATCTGTTAGGGGAGA 2142
DB 4339 ACATTTACACTGCTTCTCGTTTCAGGTCAATGATCAAGTTGTTGATTTGCTCTTGAATA 4280
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAGTTGCTTTC 2202
DB 4279 AGGAGCTCCAATCAGTGCTTAAGCAAAAAAATGTTTGGCTCCCTTACATATGCGACACA 4220
QY 2203 TCTGTATATGTAAATTTGCGGCGATACCCCAAGCAGACATTTCAACACCTTTTACATTTAGC 2262
DB 4219 GGTGGATGATGTACTGTTGACTATCTC-----ACTCCTCTTTCATGTGCGC 4175

QY 2263 AGCTGGTTATAATAATTTAGAAAGTTTGAGAGTATTTTGTACAAACACGAGCTGATGTGAA 2322
DB 4174 TGCTCATTTGGGACATGTCGTCGCTAAACATTTTCTGCTGATGTAATGCTGACCCGAA 4115
QY 2323 TGCCCAAGACAAAGAGGAGGACTTATTCCTTTTACATAAATGCAGCATCTTTACGGGCAATGAGA 2382
DB 4114 TGCTCGAGCTCTCAATGGCTTTCACACGCTGCATATCGCTTGCAAAAAAATCGCATTA 4055
QY 2383 TGTAGCAGCTCTACTATAAAGTATAATGATGTGTCATGTCAGTCCACGACAAATGGGCTTT 2442
DB 4054 AATTTGTCGAAGTCTGCTACTGAAATACCACGCTGCAATCGAAGCACTACTGAATCCGGTCT 3995
QY 2443 CACACCTTTTGACGAAGCAGCCCAAAAGAGGACGACGACGCTTTGTCTTTTGTGTAGC 2502
DB 3994 CTACCGCTGCATGCTGCTGCTTTTATGGTGTCTATAAATCTCATCTATTACTACA 3935
QY 2503 CCATGGAGTGACCCGACTCTTTAAAAATCAGGAAGACAAACACCTTTTAGATTAGTTTC 2562
DB 3934 ACAAGTGCTAATGCAGATGTGGCTACAGTACGCGTGAACGCTCTTTCATTTAGCTGC 3875
QY 2563 A 2563
DB 3874 A 3874

RESULT 9

US-08-847-429A-35
; Sequence 35, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Hesk Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5235 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-847-429A-35

Query Match 2.4%; Score 93; DB 1; Length 5235;
Best Local Similarity 45.0%; Pred. No. 1.8e-16;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

```
QY 1373 GGAGCAACATCAATGAAAGACTAAAGAAATCTTGACTCCTCTGCGACGTGGCATCTGAG 1432
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 GGCACCATATCAACACATGCAATGCGAATGGCCTTAATGCAATGCAATGCTGCCAAA 237

QY 1433 AAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGGTTAATGCTCTG 1492
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 238 GAAGGTCAATCATGAAGTGGTCCGCACTTCTGAAAGAAAGACGAGATGTTGATGCTGCC 297

QY 1493 GATAATCTTGGTCAGACTTCTTACAGAGCTGCATATTTGGTGCATCTACAAACCTGC 1552
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 298 ACTAGAAGGGTACACAGCGTTACATATAGCATCAITGGCAGGACAGAACTAATCGTC 357

QY 1553 CGCTACTCTGAGCTATGGGTGATGCTTCAACATATATACCTTTCAGGGCTTTACTGCT 1612
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 ACAGTACTTGTGAAATGTTGCTTAATGTTAACTGTAACACTAAACAGGTTTACACCA 417

QY 1613 TTACAGATGGAAATGAAATGTACAGCAACTCCTCCAGAGGGTATCTCATTAGGTTAAT 1672
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 418 CTTTACATGGCTGCACAGAAATACGAATCTGTTGTACGCTATCTTCTTGCCCAAT 477

QY 1673 TCAGAGGACAGACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGT-----CGAAAC 1725
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 478 GCCAATCAAGCTTTAAGTACAGAAAGCGGTTTACGCCACTGGCAGTTGCCTTGCACAA 537

QY 1726 TGTAAAAAACTGTGACTGTTACAGTGTCAACTGCAGAGACATTTGAAGGGGCTCAGTC 1785
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538 GGTACAGATCGTGTGTCGCTGTTTCTTGAATGACACGCGCGGAAAGTG-CGCTT 596

QY 1786 TACACCACTTCAATTTTGCAGCTGGGTAAACAGAGTCTCGTGTGGAATATCTGTACA 1845
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 597 GCAGCACTGCATATGCTGCTTAAAGAGATGATACGAAAGCAGCTACGCTATTACTTCA 656

QY 1846 GCATGGAGCTGATGTGCATCTAAAGATAAAGGAGGCTTGTACCTTTGCACATGCATG 1905
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 657 AAATGACATAAATCGGATGCTTCCGAAAGCGGCTTTACTCCGCTTCTATTCGCGC 716

QY 1906 TTCTTATGGACATATGAAGTTGCAGAACTCTTGTGTTAAACATGGAGCAGTAGTTAATG 1965
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 717 TCACTATGGAATGAGAACGTAGACACAACTCTACTCGAAAGGGGACCAATGTGAATTA 776

QY 1966 AGCTGATTTATGGAATTTACACTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAT 2025
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 777 CCAAGCAGACATACATAAGTCCGTTACACGTTGCACAAATGGGCTCGTACAAACAT 836

QY 2026 TTGCAAACTTCTGTCAGCATGGTGCAGACCCCTACCAAAAAAAGGAGGATGGAATAC 2085
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 837 GGTTCGTTATTTGTTGCTCATGGGCGCGTAATTGACTGCGCACACGTCATTTACTAAC 896

QY 2086 TCCTTTGGATCTTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 897 ACCATTACATGCTGCTCTCGTTTCAGGTGATGATCAAGTTGTTGATTTGTTGCTTGA 956

QY 2143 TGCAGCTTTTCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAGTTGTTCTTC 2202
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 957 AGGAGCTCCAATCAGTCTAAGACAAAAAATGTTTGGCTCCCTTACATATGCGAGACA 1016

QY 2203 TCCTGATAATGTAATTTGCGCGCATACCCAGGACAGACATTTCAACACCTTTTACATTTAGC 2262
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1017 GGTGGATGATGTTACTGTTGACTATCTC-----ACTCCTCTTTCATGTGCG 1061

QY 2263 AGCTGGTTATAAATTTAGAGTTTCAGAGTATTTGTTACACACAGGAGCTCATGTGAA 2322
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1062 TGTCTATGCGGAGATGTCGCTGCTGCTAACTTTTGTGATCGTAATGCTGACCGGAA 1121

QY 2323 TGCCCAAGACAAAGGAGGACTTTATCTTTTACATAATGACAGATCTTACGGCATGTAGA 2382
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1122 TGTCCAGCTCTCAATGGCTTACACCGCTGCAATATCGCTTGCACAAAAAATCGCATTA 1181

QY 2383 TGTAGAGCTCTACTAATAAGATATATGATGTCATGTCATGTCACCGACCAAAATGGGCTTT 2442
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1182 AATTGTCGAAGTCTACTGAAATACACCGTGCATCGAATCGAAGCAACTACTGAAATCCGGTCT 1241

QY 2443 CACACCTTTGCAGAGCAGCCCAAAAGGACGACACAGCTTTGTGCTTTGTGCTAGC 2502
```

```
Db 1242 CTCACCGCTGCATGCTCGTCTCTTTTATGGTGTCTATAAACATTTGCTCATCTATTACTACA 1301

QY 2503 CCATGGAGCTACCCGACTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTAGTTTC 2562
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1302 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCTCTTCTATTAGCTGC 1361

QY 2563 A 2563

Db 1362 A 1362

RESULT 10
US-08-847-429A-36/c
; Sequence 36, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5235 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-847-429A-36

Query Match 2.4%; Score 93; DB 1; Length 5235;
Best Local Similarity 45.0%; Pred. No. 1.8e-16;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;
```

```
QY 1373 GGAGCAACATCAATGAAAGACTAAAGAAATCTTGACTCCTCTGCGACGTGGCATCTGAG 1432
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5058 GGCACCGATATCAACACATGCAATGCAATGGCCTTAATGCAATTCGATTCGCTCCAAA 4959

QY 1433 AAAGCTCATATGATGTTGTTGAGTAGTGGTGAACATCAAGCAAAAGGTTAATGCTCTG 1492
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4998 GAAGGTGCATCATGAAGTGGTCCCGCACTTCTGAAAAAGAAAGCAGATGTTGATGCTGCC 4939

QY 1493 GATAATCTTGGTCAGACTTCTTACAGAGCTGCATATTTGGTGCATCTACAAACCTGC 1552
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4938 ACTAGAAGGGTAAACACAGCGTTTACATATAGCATCATTTGCCAGGACAAAGAACTAATGCTC 4879

QY 1553 CGCCTACTCTGAGCTATGGGTGTGATGCTTACATATATATCCCTTCAGGCGCTTACTGCT 1612
```



```

|||||
Db 4878 ACAGTACTTGTGAAATGGTGCTAAATGTTAACTGTAACAATCACTAAACGGTTTTACACCA 4819
QY 1613 TTACAGATGGGAATGAAATGTGACAGCAACTCTCAAGAGGTATCTCATTAGGTAAT 1672
Db 4818 CTTTACATGGGTGCACAGAAATACAGAAATCTCTGTGACGCTATCTCTTGGCCACAA 4759
QY 1673 TCAGAGGACAGACAGCAATTCGCTGAAAGCTGCAAGGCTGGAGATGT-----CGAAAC 1725
Db 4758 GCCATCAAGCTTTAAGTACAGAGACGGTTTTAGCCCACTGGCAGTTGCCCTTGCACAA 4699
QY 1726 TGTAAAAAAGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGCGCTAGTC 1785
Db 4698 GGTACGATCGTGTGCTCGCTGTTTGTCTGAAATGACACCGCGGAAAGTG-CGCTT 4640
QY 1786 TACACCACTTCATTTTGCAGCTGGGTATACAGAGTGTCCGTGGTGGAAATCTGCTACA 1845
Db 4639 GCCAGCACTGCTATATGCTGCTAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4580
QY 1846 GCATGAGCTGATGTGCATGTCTAAAGATAAAGAGGCGCTTGTACCTTGGCAATGCAATG 1905
Db 4579 AATGAGCATAACTCGGATGTGACTTCGAAAGCGGCTTTACTCGCTTCATATCGCGC 4520
QY 1906 TTTCTATGACATTAAGTGTGCAAGACTTCTTTTAAACATGGAGCAGTAAATGTT 1965
Db 4519 TCACATGAAATGAGACGTAGCACTGCTACTCGAAAGGAGCAATGTGAATTA 4460
QY 1966 AGCTGATTTATGGAATTTACACCTTTACATGAGCAGCAGCAAAAGGAAATATGAAT 2025
Db 4459 CCAAGCAGACATACATAAAGTCCGTTACAGCTTTCGCAACAAATTTGGGCTGCAACACAT 4400
QY 2026 TTGCAAACTTCTGCTCCAGCTGTGTCAGACCCCTACCAAAAAAAGAGGATGAAATAC 2085
Db 4399 GGTTCGTTATGTTGGCTCATGGGCGGTAAATGACTGTGCACACGTGAATTTACTAAC 4340
QY 2086 TCTTTTGGATCTTTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
Db 4339 ACCATTACACTGTGCTTCTCGTTTCAGGTATGATCAAGTTGTTGTTGCTTGAATA 4280
QY 2143 TGCAGCTTTGTAGATGCTGCCAAGAGGTTGTTTACCCAGAGTGAAGAAGTTGCTTTC 2202
Db 4279 AGGAGCTCCAATCAGTGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4220
QY 2203 TCTGATATGTAATTTGCGCGATACCAAGGACAGACATTCACACCTTTTACATTTAGC 2262
Db 4219 GGTGATGATGTTACTGTTGACTATCTC-----ACTCCTCTCATGTGGC 4175
QY 2263 AGCTGGTTATAATAATTTAGAGTTTGCAGAGTATTTGTTTACAACACGAGCTGATGTGAA 2322
Db 4174 TGTCTATTGCGGACATGTCGCTGTCGCTTAAACTTTTGTGCTGATGTAATGCTGACCCGAA 4115
QY 2323 TGCCCCAAGACAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGCATGTAGA 2382
Db 4114 TGCTGAGCTCTCAATGAGCTTCCACACCGCTGCATATCGCTTGCAAAAAAATTCGCAATTA 4055
QY 2383 TGTAGCAGCTCTACTAATAAAGTAAATGCAATGTCATGTCATGTCAGGACAAATGGGCTTT 2442
Db 4054 AATTGTGCAACTGCTACTCAATACCAAGCTGCAATCGAAGCACTACTGATCGGCTCT 3995
QY 2443 CACACCTTTGACAGAGACGCCAAGAGGAGCAACACAGCTTTGCTGTTTGTGCTAGC 2502
Db 3994 CTCACCGCTGCATGCTGCTGCTTTTATGGGTGCTATAAACATTTGTCATCTATTACTACA 3935
QY 2503 CCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAGCAAAACACCTTTAGATTAGTTTC 2562
Db 3934 ACAAGGTGTAATGCAGATGTGGCTACAGTACGCGGTGAAAGCGCTCTTCATTTAGCTGC 3875
QY 2563 A 2563
Db 3874 A 3874
```

RESULT 11

```

US-09-065-474-35
: Sequence 35, Application US/09065474
: Patent No. 6063599
: GENERAL INFORMATION:
: APPLICANT: Tang, Liang
: TITLE OF INVENTION: DIOPHILARIA AND BRUGIA ANKYRIN
: TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
: TITLE OF INVENTION: USES THEREOF
: NUMBER OF SEQUENCES: 171
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Carol Talkington Verser, Ph.D.
: ADDRESSEE: Heska Corporation
: STREET: 1825 Sharp Point Drive
: CITY: Fort Collins
: STATE: Colorado
: COUNTRY: USA
: ZIP: 80525
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: Windows 95
: SOFTWARE: WordPerfect for Windows, Version 7.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/065,474
: FILING DATE: 24-APR-1998
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Verser, Carol Talkington
: REGISTRATION NUMBER: 37,459
: REFERENCE/DOCKET NUMBER: HW-5-Q1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 970/493-7272
: TELEFAX: 970/484-9505
: INFORMATION FOR SEQ ID NO: 35:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5235 nucleotides
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-09-065-474-35
```

```

Query Match 2.4%; Score 93; DB 3; Length 5235;
Best Local Similarity 45.0%; Pred. No. 1.8e-16;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

QY 1373 GGAGCAACATCAATGAAAAAGACTTAAAGAAATTTTGACTCCTCTGACGCTGGCATCTGAG 1432
Db 178 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCAITTGCAITGGCTCCAAA 237
QY 1433 AAAGCTCATATGATGTTGTGAAGTAGTGTGAACATGAACAAAGGTTAATGCTCTG 1492
Db 238 GAAGGTCAATGATGAAGTGGTCCGCGAATCTTCTGAAAGAAAAGACGATGTTGATGCTGC 297
QY 1493 GATAATCTTGCTAGACTTCTCTACACAGAGCTGCATATTGTGCTCATCTACAACCTGC 1552
Db 298 ACTAGAAGGGTACACAGCCTTACATATAGCATATTGGCAGGACAAGAACTAATCGTC 357
QY 1553 CGCCTACTCTGAGCTATGGGTGTGATCCTCAACATTATATCCCTTTCAGGGCTTTACTGCT 1612
Db 358 ACAGTACTTGTGAAAAATGGTCTTAATGTTAAACGTACACTCACTTAAACGGTTTTTACACCA 417
QY 1613 TTACAGATGGGAATGAAATGTACAGCAACTCCTCCAAGGGGTATCTCATTAGGTAAT 1672
Db 418 CTTTACATGGCTGCAGAGAAAATCAGCAATCTGTTGTACGCTATCTTCTTGGCCCAAT 477
QY 1673 TCAGAGGACAGACAAATTCCTGGAAGCTGCAAGGCTGGAGATGT-----CGAAAC 1725
Db 478 GCCAATCAAGCTTTAAGTACAGAGACGGTTTTTACGCCACTGGCAGTTGCTTGCACAA 537
QY 1726 TGTAAAAAAGTGTGTTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGGTGCTAGTC 1785
Db 1726 TGTAAAAAAGTGTGTTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGGTGCTAGTC 1785
```

Db 538 GGTCCAGATCGTGGTGGCTGTTTCTTGAATAAGACACGCGCGGAAAGTG-CGCTT 596
QY 1786 TACACCATTCTTTTCAGCTGGGTATACAGAGTGTCCCTGGTGAATATCTGTACA 1845
Db 597 GCCAGCACTGCATATCTCTAAAGAGATGATACAAAGCAGCTACGCTATTACTTCA 656
QY 1846 GCATGGAGCTGATGTCATGCTAAAGATAAAGGAGGCTTGTACCTTTGGCACAATGATG 1905
Db 657 AAATGAGCATAAATCGCATGTGACTTCCGAAAGCGGCTTACTCCGCTTCATATCGCCGC 716
QY 1906 TTCTTATGACATATCAAGTTCAGACACTTCTTGTAAACATGGACGAGTAGTAATGT 1965
Db 717 TCACATGGAATGAGAACGTGACACACTGCTACTCGAAAGGAGCGCAATGTGAATTA 776
QY 1966 AGCTGATTTATGGAATTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAT 2025
Db 777 CCAAGCAGACATAACATAAGTCCGTTACAGTTGCAACAAATGGGCTGCTACAACAT 836
QY 2026 TTGCAAACTTCTGTCAGCATGGTGCAGACCCCTACCAAAAAAACAGGATGGAATAC 2085
Db 837 GGTTCCTTATTTGGCTCATGGGCGGTAATGTACTGTCGCACACCTGATTTACTAAC 896
QY 2086 TCCTTTGATCTTGTAAAGATGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
Db 897 ACATTTACATGCTCTCTCTCAGTGCATGATCAAGTTGTTGATTTGTTGCTTGAATA 956
QY 2143 TGCAGCTTTTCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGAGTGAAGAGTTGTCTTC 2202
Db 957 AGGAGTCCCAATCAGTCTAAGACAAAATAATGTTTGGCTCCCTTACATATGCGACACA 1016
QY 2203 TCCTGATAATGTAANTTGGCGGATACCCAAAGGAGCAGACATTTACACCTTTTACATTTAGC 2262
Db 1017 GGTGGATGATGTTACTGTTGACTATCTC-----ACTCCTCTTCATGTGGC 1061
QY 2263 AGCTGTTTATAATAATTAGAACTTGCAGAGTATTTGTTTACACACGAGCTGATGTAA 2322
Db 1062 TGCTCATGCGGACATGTCGCTGCGTAACTTTTGTGGATCGTAATGCTGACCGGAA 1121
QY 2323 TGCCCAAGACAAAGGAGGACTTATCTTTACATAATGACAGCATTTACGGGATGTAGA 2382
Db 1122 TGCTCAGCTCTCAATGGCTTTCACACCGTGCATATCGCTTGCAAAAAAATCGCATTA 1181
QY 2383 TGTCAGAGCTCTACTAATAAGTATATGATGTGTCATGTCATGTCACGCAAAATGGCGTTT 2442
Db 1182 AATTGCGAACTGCTACTGAAATACACCGCTGCAATTCGAAGCAACTACTGAATCCGCTCT 1241
QY 2443 CACACCTTTGACGAGAGCCCAAGGAGCAACACAGCTTTGTGCTTTGTGTGCTAGC 2502
Db 1242 CTCACCGCTCATGTCGCTGCTTTTATGGTGCTATAAACAATGTCATCTATTACTACA 1301
QY 2503 CCATGAGCTGACCCGACTTTAAAATCAGGAAGGACAAACACCTTTTAGATTTAGTTTC 2562
Db 1302 ACAAGTGCTAATGCAATGTGGCTACAGTACGCGGTGAACAGCCCTCTTCATTAGCTGC 1361
QY 2563 A 2563
Db 1362 A 1362

RESULT 12

US-09-065-474-36/c

; Sequence 36, Application US/09065474

; Patent No. 6063599

; GENERAL INFORMATION:

; APPLICANT: Tang, Liang

; APPLICANT: Blehm, E. Scot

; TITLE OF INVENTION: DIFOLIARIA AND BRUGIA ANKYRIN

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND

; TITLE OF INVENTION: USES THEREOF

; NUMBER OF SEQUENCES: 171

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carol Talkington Verser, Ph.D.

; ADDRESSEE: Heska Corporation

; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/494-9505
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5235 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-065-474-36

Query Match 2.4%; Score 93; DB 3; Length 5235;

Best Local Similarity 45.0%; Pred. No. 1.8e-16;

Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

QY 1373 GGACCAACATCAATGAAAGACCTAAAGAAATCTTGACTCTCTGACAGTGGCATCTGAG 1432

Db 5058 GGACCGATATCAACACATGCAATGCAATGGCTTAAATGCAATGCTGGCTCCAAA 4999

QY 1433 AAAGCTCAATATGATGTTGAAGTAGTGTAACATGAACAAAGTTAATGCTCTG 1492

Db 4998 GAAGTCAATCAAGTGTCCCGGAATCTTGAAGAAGAAAGAGATGTTGATGCTGCC 4939

QY 1493 GATAATCTTGGTCAGACTTCTCTACAGAGCTGCATATTTGTGTCTATCTACAAACCTGC 1552

Db 4938 ACATGAAAGGTAACACAGGTTACATATAGCATCATTTGCGAGCAAGAATTAATCGTC 4879

QY 1553 CGCTACTCTGAGCTATGGTGTGATCTTAAATATATATCCCTTCAGGCGTTTACTGCT 1612

Db 4878 ACAGTACTTGTGAAATGGTGTAAATGTTAAGTACAAATCACTAAACGGTTTTTACACCA 4819

QY 1613 TTACAGATGGAAATGAAATGTACAGCAACTCTCCAGAGGGTATCTCATTTAGGTAAT 1672

Db 4818 CTTTACATGGCTGCACAGAAATCAGAAATCTGTTGACGCTATCTTCTTGGCCACAAT 4759

QY 1673 TCAGAGGACAGACAATTTGCTGGAAGCTGCAAGGCTGGAGATGT-----CGAAAC 1725

Db 4758 GCAATCAAGCTTTAAGTACAGAAAGACGGTTTTACGCCACTGCGCTTGGCTTGAACAA 4699

QY 1726 TGTAAGAAACCTGTGTACTGTTCCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTC 1785

Db 4698 GGTACAGATCGTGTGCTGCTGTTTGTGTTGAAATGACACGCGCGGAAAGTG-CGCTT 4640

QY 1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTTGGTGAATATCTGTACA 1845

Db 4639 GCCAGCACTGCATATTTGCTGCTAAAAGAGATGATAGAAAGCAGCTACGCTATTACTTCA 4580

QY 1846 GCATGAGCTGATGTGCATGCTTAAAGATAAGAGGCGCTTGTACCTTTGCACAATCATG 1905

Db 4579 AAATGAGCATAACTCGGATGTGACTTCGAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4520

QY 1906 TTCTTATGACATATGAAAGTTCGCAACTTCTTGTAAACATGAGCAGTAGTAAATGT 1965

Db 4519 TCACATGGAATGAAAGCTAGCACTGCTACTCGAAAGGAGGCAATGTGAATTA 4460

QY 1966 AGCTGATTTATGGAATTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAAT 2025
Db 4459 CCAAGCGACACATACATAAGTCCGTTACAGTTGCACAAAATGGGTGCTGACAAACAT 4400
QY 2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCTACCAAAAAACAGGGATGGAATAC 2085
Db 4399 GGTTCGTTATTTGGTCTGAGCGGCGTAAATGACTGTGCGACACGTGATTTACTAAC 4340
QY 2086 TCCTTTGGATCTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
Db 4339 ACAATTACACTGTGCTTCGTTACAGGTCATGATCAAGTTGTTGATTTGCTTGAATA 4280
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTACGAGAGTGAAGAAGTTGCTTTC 2202
Db 4279 AGGAGCTCAATCAGTGTACACAAAATAAGTTTGGCTCCCTTACATATGGCAGCACA 4220
QY 2203 TCCTGATAATGTAATTTGCCGGATACCCAGGCGACACATTAACACCTTTTACATTTAGC 2262
Db 4219 GGTGGATGATGTTACTGTTGACTATC-----ACTCCTCTTCATGTGCG 4175
QY 2263 AGCTGCTTATAATTTAGAACGTTGACAGATATTTTACACACAGGAGCTGATGTGAA 2322
Db 4174 TGTCTAATCGGACATGTCGCTGCGTAAACTTTTGTGGATCGTAAATGCTGACCCGAA 4115
QY 2323 TGCCCAAGACAAAAGGAGGACTTATTCCTTTACATAATCGAGCATCTTACGGGCATGTAGA 2382
Db 4114 TGCTCGAGCTCTCAATGGCTTACACCGCTGCATATCGTTGCCAAAAAATCGCATTA 4055
QY 2383 TGTTAGCAGCTCTACTAATAAAGATATAATGATGTGTCAATGCCAGGACAAATGGGCTTT 2442
Db 4054 AATTGTCGAACCTGCTACTGAAATACACAGCTGCAATCGAAGCAACTACTGAAATCGGTCT 3995
QY 2443 CACACCTTTGCAGGAGCAGCCCAAGAGGAGGACACAGCTTTGTGCTTGTCTAGC 2502
Db 3994 CTCACGCTGCTATGCTGCTGCTTTTATGGGTGCTATAAACAATGTCATCTATTACTACA 3935
QY 2503 CCATGAGCTGATCCCGACCTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTAGTTTC 2562
Db 3934 ACAAGTGCTAATGCAGATGTGCTACAGTACGCGGTGAAGCGCTCTTCATTTAGCTGC 3875
QY 2563 A 2563
Db 3874 A 3874

RESULT 13
US-09-557-034-35
; Sequence 35, Application US/09557034
; Patent No. 6365569
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; PROTEINS, NUCLEIC ACID MOLECULES, AND
; USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,034
; FILING DATE: 21-Apr-2000

; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065,474
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5235 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-557-034-35

Query Match 2.4%; Score 93; DB 4; Length 5235;
Best local Similarity 45.0%; Pred. No. 1.8e-16;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;
QY 1373 GGAGCAAAACATCAATGAAAGAGCTAAAGAAATTTCTTGACTCTCTGCACGTGGCATCTGAG 1432
Db 178 GGCACCGATATCAACACATGCAATGCGAATGGCCCTTAATGATTGCATCTGGCCCTCAAA 237
QY 1433 AAAGCTCATATGATGTTGTGAAGTAGTGTGAACATGAAGCAAGAGTTAATGCTCTG 1492
Db 238 GAAGTTCATCATGAAGTGGTCCGCGAACTTCTGAAAAAGAAAGCAGATGTTGATGCTGCC 297
QY 1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGTGTCATCTACAAACCTGC 1552
Db 298 ACTAGAAGGGTACACAGCGCTTACATATATGATCATTTGGCAGGACAAAGCACTAATCGTC 357
QY 1553 CGCTACTCTCTGAGCTATGGGTGTGATCCCTAACATTTATATCCCTTCAGGGCTTTACTGCT 1612
Db 358 ACAGTACTTGTGAAATGGTCTAATGTTAACGTACAACTACCTAAAGGCTTTTACACCA 417
QY 1613 TTACAGATGGGAAATGAAATGTACAGCAACTCTCCAAGAGGGTATCTCATTAAGTAAAT 1672
Db 418 CTTTACATGGCTGCACAAAGAAATCAGCAATCTGTTGTAGCTATCTTCTTGCCCAAT 477
QY 1673 TCAGAGGACAGACACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGT-----CGAAC 1725
Db 478 GCCAATCAAGCTTTAAGTACAGAAAGAGCGGTTTACGCCACTGGCAGTTGCTTTGCAACAA 537
QY 1726 TGTAATAAACTGTGTACTGTTACAGAGTGTCAACTGTCAGAGACATTTGAAGGGCTCAGTC 1785
Db 538 GGTACAGATCGTGTGCTGCTGCTTTTCTGTTGAAATACACCGCGCGGAAAGTG-CGCTT 596
QY 1786 TACACCCTTCAATTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCTACA 1845
Db 597 GCCAGCACTGCATATGCTGCTAAAAAAGATGATCAAGAAAGCAGCTACGCTATTACTTCA 656
QY 1846 GCATGGAGCTGATGTGATGCTAAAGATAAAGGAGGCTTGTACCTTTGCACATGCTATG 1905
Db 657 AAATGAGCATAACTCGGATGTGACTTTCGAAAGCGGCTTTACTCCGCTTATATCGCCGC 716
QY 1906 TTCTTATGGACATTTAAGTGTGAGAGTCTCTTGTGTTAAACATGGAGCAGTAGTAAATGT 1965
Db 717 TCATATGGAATGAGACGTAGCACACTGCTACTCGAAAGGAGGCAATGGAATTA 776
QY 1966 AGCTGATTTATGGAATTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAAT 2025
Db 777 CCAAGCGAGACATAACATAAGTCCGTTACAGCTTTCACAAATAATGGGTGCTGCAACACAT 836
QY 2026 TTGCAAACTTCTGCTCCAGCATGTTGTCAGACCCCTTACCAAAAAACAGGATGGAATAC 2085
Db 837 GGTTCGTTATTTGTTGGCTCATGGGCGGTAATTTGACTGTGCGACACGTGATTACTTAAC 896

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 18:26:28 ; Search time 167.581 Seconds
(without alignments)
11540.122 Million cell updates/sec

Title: US-09-843-159B-1

Perfect score: 3797

Sequence: 1 cttgaagacactggatttc.....cctttgctgaaaaaaaaa 3797

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3297	86.8	3400	10 US-09-509-196A-1	Sequence 1, Appli
2	1813.4	47.8	4134	10 US-09-841-835-1	Sequence 1, Appli
3	1449.6	38.2	4491	10 US-09-841-835-7	Sequence 7, Appli
4	1297.8	34.2	4657	10 US-09-841-835-9	Sequence 9, Appli
5	994	26.2	2409	9 US-09-964-899-40	Sequence 40, Appli
6	467.6	12.3	1069	10 US-09-833-381-841	Sequence 841, App
7	461	12.1	465	10 US-09-833-381-1153	Sequence 1153, Ap
8	323	8.5	353	9 US-10-040-739-1179	Sequence 1179, Ap
9	121.2	3.2	523	10 US-09-841-835-12	Sequence 12, Appli
10	117.6	3.1	304	10 US-09-783-590-10862	Sequence 10862, A
11	81.6	2.1	466	10 US-09-833-381-842	Sequence 842, App
12	80.8	2.1	5175	9 US-09-964-899-42	Sequence 42, Appli
13	77	2.0	1299	10 US-09-908-711-12	Sequence 12, Appli
14	74.2	2.0	2505	10 US-09-947-199-3	Sequence 3, Appli
15	74.2	2.0	3025	10 US-09-947-199-1	Sequence 1, Appli
16	73.6	1.9	948	9 US-09-938-842A-2359	Sequence 2359, Ap
17	73.2	1.9	2271	10 US-09-835-788A-2	Sequence 2, Appli
18	72.4	1.9	744	9 US-09-938-842A-805	Sequence 805, App
19	69.6	1.8	357	10 US-09-841-835-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-09-509-196A-1

; Sequence 1, Application US/09509196A

; Patent No. US20020037582A1

; GENERAL INFORMATION:

; APPLICANT: DALY, Roger J.

; TITLE OF INVENTION: A Potential Effector for the Grb7 Family of Signalling

; TITLE OF INVENTION: Proteins

; FILE REFERENCE: 1871-129

; CURRENT APPLICATION NUMBER: US/09/509,196A

; CURRENT FILING DATE: 2000-03-23

; PRIOR APPLICATION NUMBER: P09388

; PRIOR FILING DATE: 1997-09-23

; PRIOR APPLICATION NUMBER: PCT AU98/00795

; PRIOR FILING DATE: 1998-09-23

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 3400

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-509-196A-1

Query Match 86.8% Score 3297; DB 10; Length 3400;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 3303; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 482 ATTCTCTTCATAATGCGATGCTCTTTGGTCATGCTGAAGTAGTCAATCTCTTTTGGCA 541

|||||

Db 1 ATTCCTCTTCATAATGCGATGCTCTTTGGTCATGCTGAAGTAGTCAATCTCTTTTGGCA 60

Qy 542 CATGGTCGAGACCCCAATGCTCGAGATAATGGAAATATATCTCTTCCATGAAGCTGCA 601

|||||

Db 61 CATGGTCGAGACCCCAATGCTCGAGATAATGGAAATATATCTCTTCCATGAAGCTGCA 120

Qy 602 ATTAAGAAAGATTGATGTTTGCAATGCTGTGTACAGCATGCGCTGAGCCCAACATC 661

|||||

Db 121 ATTAAGAAAGATTGATGTTTGCAATGCTGTGTACAGCATGCGCTGAGCCCAACATC 180

Qy 662 CGAAATACAGATGAAGGACAGCATTTGATTTAGCAGATCCATCTGCCAAGAGAGTGCTT 721

|||||

Sequence 253, App
Sequence 1141, App
Sequence 310, App
Sequence 7, Appli
Sequence 1406, App
Sequence 1406, App
Sequence 151, App
Sequence 1150, App
Sequence 9, Appli
Sequence 7, Appli
Sequence 1301, App
Sequence 3, Appli
Sequence 182, App
Sequence 57, Appli
Sequence 568, App
Sequence 568, App
Sequence 29, Appli
Sequence 1149, App
Sequence 443, App
Sequence 2, Appli
Sequence 48, Appli
Sequence 2, Appli
Sequence 134, App
Sequence 5, Appli
Sequence 1151, App

Db 181 CGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAAGCAGTCTT 240
QY 722 ACTGGTGAATATAGAAGATGAACCTCTAGAAAGTGCAGGAGTGGCAATGAAGAAAAA 781
Db 241 ACTGGTGAATATAGAAGATGAACCTCTAGAAAGTGCAGGAGTGGCAATGAAGAAAAA 300
QY 782 ATGATGGCTCTACTCACACCATTTAAATGTCACTGCCACGAAGTGATGSCAGAAAGTCA 841
Db 301 ATGATGGCTCTACTCACACCATTTAAATGTCACTGCCACGAAGTGATGSCAGAAAGTCA 360
QY 842 ACTCCATTACATTTGGCAGCAGATATACAGAGTAAGAATTTGACAGCTGTTACTGCAA 901
Db 361 ACTCCATTACATTTGGCAGCAGATATACAGAGTAAGAATTTGACAGCTGTTACTGCAA 420
QY 902 CATGGAGCTGATGCCATGCTAAAGATAAAGTGATCTGGTACCAATTACACAATGCTGT 961
Db 421 CATGGAGCTGATGCCATGCTAAAGATAAAGTGATCTGGTACCAATTACACAATGCTGT 480
QY 962 TCTTATGGTCAATATGAAGTAACAGTGAACCTTTGGTCAAGCATGGTGCCTGTGTAATGCA 1021
Db 481 TCTTATGGTCAATATGAAGTAACAGTGAACCTTTGGTCAAGCATGGTGCCTGTGTAATGCA 540
QY 1022 ATGGACTTGTGGCAATTCACCTCTCTCATGAGCGAGCTCTTAAGACAGGGTTGAAGTA 1081
Db 541 ATGGACTTGTGGCAATTCACCTCTCTCATGAGCGAGCTCTTAAGACAGGGTTGAAGTA 600
QY 1082 TGTTCCTCTCTTAAAGTTATGGTGAGACCCCAACACTGCTCAATTTGTACAAATAAAGT 1141
Db 601 TGTTCCTCTCTTAAAGTTATGGTGAGACCCCAACACTGCTCAATTTGTAAAGATAAAGT 660
QY 1142 GCTATAGACTTGGCTCCACACACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGC 1201
Db 661 GCTATAGACTTGGCTCCACACACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGC 720
QY 1202 CACTCGTGTGCAAGTGCACGAGAGCTGATGTTACTCGAATCAAAAACATCTCTCT 1261
Db 721 CACTCGTGTGCAAGTGCACGAGAGCTGATGTTACTCGAATCAAAAACATCTCTCT 780
QY 1262 CTGGAATAGTGTAATTTCAAGCATCCTCAACACATGAACAGCATTTGCATTTGCTGCT 1321
Db 781 CTGGAATAGTGTAATTTCAAGCATCCTCAACACATGAACAGCATTTGCATTTGCTGCT 840
QY 1322 GCATCTCCATATCCAAAAGAAAGCAAAATATGTGAATGTTGCTAAGAAAGAGCAAAAC 1381
Db 841 GCATCTCCATATCCAAAAGAAAGCAAAATATGTGAATGTTGCTAAGAAAGAGCAAAAC 900
QY 1382 ATCAATGAAAGACTAAAGAAATTTCTTGACTCTTCACGTGSCATCTGAGAAAGCTCAT 1441
Db 901 ATCAATGAAAGACTAAAGAAATTTCTTGACTCTTCACGTGSCATCTGAGAAAGCTCAT 960
QY 1442 AATGATGTTGTGAAGTAGTGGTGAACACATGAAGCAAGGTTAATGCTCTGGATAATCTT 1501
Db 961 AATGATGTTGTGAAGTAGTGGTGAACACATGAAGCAAGGTTAATGCTCTGGATAATCTT 1020
QY 1502 GGTACACATCTCTACAGAGCTGCATATTTGGTGTCATCTACAAACCTGCGCCCTACTC 1561
Db 1021 GGTACACATCTCTACAGAGCTGCATATTTGGTGTCATCTACAAACCTGCGCCCTACTC 1080
QY 1562 CTGAGCTATGGGTGATCCTTAACATATATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1621
Db 1081 CTGAGCTATGGGTGATCCTTAACATATATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140
QY 1622 GGAATGAAATGTACAGCAACTCTCCAGAGGGTATCTCATTTAGGTAAATTCAGAGGCA 1681
Db 1141 GGAATGAAATGTACAGCAACTCTCCAGAGGGTATCTCATTTAGGTAAATTCAGAGGCA 1200
QY 1682 GACAGACAATTTGCTGGAAGTGCAGAGGCTGGAGATGTGAAACTGTAAAGAACTGTGT 1741
Db 1201 GACAGACAATTTGCTGGAAGTGCAGAGGCTGGAGATGTGAAACTGTAAAGAACTGTGT 1260
QY 1742 ACTGTTTCAGAGTCTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACCCTTCATT 1801
Db 1261 ACTGTTTCAGAGTCTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACCCTTCATT 1320

QY 1802 GCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGTACAGCATGGAGCTGATGTG 1861
Db 1321 GCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGTACAGCATGGAGCTGATGTG 1380
QY 1862 CATGCTAAAGATAAAGAGGCCCTTGACCTTTGCACAAATGCATGTTCTTTATGACATAT 1921
Db 1381 CATGCTAAAGATAAAGAGGCCCTTGACCTTTGCACAAATGCATGTTCTTTACGGACATAT 1440
QY 1922 GAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGAAA 1981
Db 1441 GAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGAAA 1500
QY 1982 TTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTTCTGCTC 2041
Db 1501 TTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTTCTGCTC 1560
QY 2042 CAGCATGGTGAGACCTTACCACCAAAACAGGAGTGGAAATACTCTTTTGGATCTTTGTT 2101
Db 1561 CAGCATGGTGAGACCTTACCACCAAAACAGGAGTGGAAATACTCTTTTGGATCTTTGTT 1620
QY 2102 AAAGATGGAGATACAGATATTCAGAGTCTGCTTAGGGAGATGCAGCTTTGCTAGATGCT 2161
Db 1621 AAAGATGGAGATACAGATATTCAGAGTCTGCTTAGGGAGATGCAGCTTTGCTAGATGCT 1680
QY 2162 GCCAAGAAGGGTGTGTAGCCAGAGTGAAGATTTGTTCTCTCTGATATGTAATTTG 2221
Db 1681 GCCAAGAAGGGTGTGTAGCCAGAGTGAAGATTTGTTCTCTGATATGTAATTTG 1740
QY 2222 CGGGATACCCAGGAGACATTCACACCTTTACATTTAGCAGCTGGTTATATATTTA 2281
Db 1741 CGGGATACCCAGGAGACATTCACACCTTTACATTTAGCAGCTGGTTATATATTTA 1800
QY 2282 GAAGTTGCAGAGTATTTGTACAAACAGGAGCTGTGAATGCCCAAGCAAGAGGAGGA 2341
Db 1801 GAAGTTGCAGAGTATTTGTACAAACAGGAGCTGTGAATGCCCAAGCAAGAGGAGGA 1860
QY 2342 CTTTATTCCTTTTACATAATGCAGCATCTTACGGCATGTAGATGTAGCAGCTCTACTAATA 2401
Db 1861 CTTTATTCCTTTTACATAATGCAGCATCTTACGGCATGTAGATGTAGCAGCTCTACTAATA 1920
QY 2402 AAGTATATGTCATGTGTAATGCCAGGACAAATGGCTTTTCACACCTTTTCACAGAGCA 2461
Db 1921 AAGTATATGTCATGTGTAATGCCAGGACAAATGGCTTTTCACACCTTTTCACAGAGCA 1980
QY 2462 GCCCAAAAGGAGCAACACAGCTTTTGTGCTTGTAGCCCATGGAGCTGACCCGACT 2521
Db 1981 GCCCAAAAGGAGCAACACAGCTTTTGTGCTTGTAGCCCATGGAGCTGACCCGACT 2040
QY 2522 CTTTAAAAATCAGGAAGCAAAACACCTTTTAGATTTAGTTTTCAGCGGATGATGTACGCGCT 2581
Db 2041 CTTTAAAAATCAGGAAGCAAAACACCTTTTAGATTTAGTTTTCAGCGGATGATGTACGCGCT 2100
QY 2582 CTTTCTCAGCAGCCCATGCCCCCATCTGCTCTGCCCTCTTGTGTACAGCCTCAAGTGCTC 2641
Db 2101 CTTTCTCAGCAGCCCATGCCCCCATCTGCTCTGCCCTCTTGTGTACAGCCTCAAGTGCTC 2160
QY 2642 AATGTTGTGAGAGCCAGGAGCCTTCGAGATGCTCTCTCTTCAGCTCCATCTAGCCCA 2701
Db 2161 AATGTTGTGAGAGCCAGGAGCCTTCGAGATGCTCTCTCTTCAGCTCCATCTAGCCCA 2220
QY 2702 TCAAGCCTTTCTCAGCCAGCAGCTTTGACAACCTTATCTGGGAGTTTTTCAGAACTGTCT 2761
Db 2221 TCAAGCCTTTCTCAGCCAGCAGCTTTGACAACCTTATCTGGGAGTTTTTCAGAACTGTCT 2280
QY 2762 TCAGTAGTTAGTTCAAGTGAACAGAGGCTTTCAGTTTGGAGGAGTTTTTCAGAACTGTCT 2821
Db 2281 TCAGTAGTTAGTTCAAGTGAACAGAGGCTTTCAGTTTGGAGGAGTTTTTCAGAACTGTCT 2340
QY 2822 GGAAGTAGTTTTCAGTAACTCAATTCGTAAGGAATCTTGGACTTTCAGCAGCTTAAATGAT 2881
Db 2341 GGAAGTAGTTTTCAGTAACTCAATTCGTAAGGAATCTTGGACTTTCAGCAGCTTAAATGAT 2400

QY	2882	ATATTTGAGAGAGACAGATCACTTTGGATGATTAGTTGAGATGGGGCACAAGGAGCTG	2941
Db	2401	ATATTTGAGAGAGAACAGATCACTTTGGATGATTAGTTGAGATGGGCAACAAGGAGCTG	2460
QY	2942	AAGCAGATTGGAACTCAATGCTTATGCACATAGGCACAACACTAATTTAAGGAGACTCGACAGA	3001
Db	2461	AAGGAGATTGGAACTCAATGCTTATGCACATAGGCACAACACTAATTTAAGGAGACTCGACAGA	2520
QY	3002	CTTATCTCCGGACAACAAGGCTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGGA	3061
Db	2521	CTTATCTCCGGACAACAAGGCTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGGA	2580
QY	3062	ACAATTCCTTATAGATCTGCTCTCGTAGTATAAGAGTTTCAGTCTGTGGAGGAAGAGATG	3121
Db	2581	ACAATTCCTTATAGATCTGCTCTCGTAGTATAAGAGTTTCAGTCTGTGGAGGAAGAGATG	2640
QY	3122	CAAAAGTACAGTTCGAGAGCAGACAGATGAGGTCATGCAGGTGGAACTCTTCAACAGATAC	3181
Db	2641	CAAAAGTACAGTTCGAGAGCAGACAGATGAGGTCATGCAGGTGGAACTCTTCAACAGATAC	2700
QY	3182	AATATTCCTCAAGATTCAAGAGTTTGTACAAAGAACTATATGGGAAAGATACACTCACCGG	3241
Db	2701	AATATTCCTCAAGATTCAAGAGTTTGTACAAAGAACTATATGGGAAAGATACACTCACCGG	2760
QY	3242	AGAAAGAAGTTTCTGGAAGAAACCAACACCATGCCAATGAACGAATGCTATTTTCATGGG	3301
Db	2761	AGAAAGAAGTTTCTGGAAGAAACCAACACCATGCCAATGAACGAATGCTATTTTCATGGG	2820
QY	3302	TCGCCCTTTGCTGAATGCAATTATCCACAAGAGCTTTGATGAAAGGCGATGGCTACATAGGT	3361
Db	2821	TCGCCCTTTGCTGAATGCAATTATCCACAAGGCTTTGATGAAAGGCGATGGCTACATAGGT	2880
QY	3362	GGTATGTTTGGAGCTGGCATTTATTTTGGCTGAAAACTCTTCCAAAAGCAATCAATATGTA	3421
Db	2881	GGTATGTTTGGAGCTGGCATTTATTTTGGCTGAAAACTCTTCCAAAAGCAATCAATATGTA	2940
QY	3422	TATGGAATTGGAGGAGGTACTGGGTGTCAGTTCACAAGACAGATCTTGTTCACATTTGC	3481
Db	2941	TATGGAATTGGAGGAGGTACTGGGTGTCAGTTCACAAGACAGATCTTGTTCACATTTGC	3000
QY	3482	CACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCTTTCTGCGAGTTTCAGTGCA	3541
Db	3001	CACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCTTTCTGCGAGTTTCAGTGCA	3060
QY	3542	ATGAAATGGCAGATTCTCCTCCAGGTCACTCACTCAGTCAGTGTAGGCCAGCTGTAAT	3601
Db	3061	ATGAAATGGCAGATTCTCCTCCAGGTCACTCACTCAGTCAGTGTAGGCCAGCTGTAAT	3120
QY	3602	GGCCTAGCATTTAGCTGAATATGTTATTTTACAGAGAGAACAGGCTTATCCTGAGTATTTA	3661
Db	3121	GGCCTAGCATTTAGCTGAATATGTTATTTTACAGAGAGAACAGGCTTATCCTGAGTATTTA	3180
QY	3662	ATTACTTACCAGATTATGAGGCCCTGAAAGGTATGTTGCGATGGATAAATAGTATTTTAAGA	3721
Db	3181	ATTACTTACCAGATTATGAGGCCCTGAAAGGTATGTTGCGATGGATAAATAGTATTTTAAGA	3240
QY	3722	AACTAATTCCTAGTGAACCTTAAATCATCAAAAGCAGCAGTGGCCCTACGTTTACTCCCTT	3781
Db	3241	AACTAATTCCTAGTGAACCTTAAATCATCAAAAGCAGCAGTGGCCCTACGTTTACTCCCTT	3300
QY	3782	TGCTGAAAAAAA 3794	
Db	3301	TGCTGAAAAAAA 3313	

RESULT 2

```

US-09-841-835-1
; Sequence 1, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS

```

Db 1077 CTAATGTGAATGCCATGCAAGTGAATGGCGAAGATCGCACTCCTTTACATCTAGCAGCG 11136
QY 863 GGATATAACAGAGTAAGATTGTACAGCTGTTACTCGACATCGAGCTGATGTCCATGCT 922
Db 1137 GGCTACAACAGAGTTGGAATGATTCAGCTCTCTCTTCAGCATGGTGTGATGTTCAATGCA 11196
QY 923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCCTGTTCTTATGGTCATTTAAGTA 982
Db 1197 AAAGACAAAGGTGCATCTGGCCCTTCATATGATGTTTCATATGGACATATTAAGTC 1256
QY 983 ACTGAATCTTTGGTCAAGCATGGTGGCTGTGTAATGCAATGCAATGAGCTGTGGCAATTCAT 1042
Db 1257 ACAGAATGCTACTAAAGCATGAGCTTGTGTTAATGCCATGATCTCTGGCAGTTTACT 1316
QY 1043 CTTCTTCATGAGCGCTCTTAAGAACAGGGTTGAAGTATGTTCTCTCTCTTTAAGTTAT 1102
Db 1317 CCACTGACGAGGCTGCTTCCAGAACCGTGTAGAAGTCTGCTCTTGTGTTACTTAGCCAT 1376
QY 1103 GGTGCGAGCCCAACACTGCTCAATTTGTACAATAAAGTGTCTATAGACTTTGGTCCACACA 1162
Db 1377 GGCCTGATCCTACGTTAGTCAACTGCGATGCGCAAAAGTCTGTGGATATGGCTCCAAC 1436
QY 1163 CCACAGTTAAACAAAGATTAGCATATGAATTTAAAGGCCACTGCTGTGCGAAGTCCA 1222
Db 1437 CCGAGCTTAGGGAGAGATTTGACTTATGAATTTAAAGGTCAITTTTACTACAAAGCGC 1496
QY 1223 CGAAGCTGATGTTACTCGAATCAAAAACATCTCTCTCTGAAATGTTGAATTTCAAG 1282
Db 1497 AGAAGCAGACACTTAGCTAAGTTAAAAAACACTCGCTCGAATCATTAATTTCAA 1556
QY 1283 CATCTCAACACATGAACAGCATTTGCAATTTGCTGTCTGCATCTCCATATCCCAAGA 1342
Db 1557 CAACCGCAGTCTCATGAACAGCACTGCATGTGCTGTGGCTCTCTGCACTCCCAACGT 1616
QY 1343 AAGCAATATGTGAATGTTGCTTAAGAAAGAGGCAACATCAATGAAGAGCTTAAGAA 1402
Db 1617 AAACAAGTGAAGATTTGTTACTTTGAAGAGGAGCAATGTTAATGAAGAAATAAAGAT 1676
QY 1403 TTTCTGACTCTCTGCACTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTG 1462
Db 1677 TTTATGACTCCCTGCACTTTGAGCGCGAAAGACCCATATGATGTCATGGAAGTTCTG 1736
QY 1463 GTGAACATGAAGCAAGGTTAATGCTGCTGGAATATCTGCTCAGACTTCTCTACACAGA 1522
Db 1737 CATGAAGATGGCGCAAGATGAATGCACTGGACACCTTGGTCAGACTGCTTTGCATAGA 1796
QY 1523 GCTGCATATTTGCTCATCTACAAACCTCGCGCTACTCTGAGCTATGGGTGATGCTCT 1582
Db 1797 GCGCCCTAGCAGGCCACCTGCGAGACCTGCGCTCTCTGCTGAGTACGGCTCTGACCCC 1856
QY 1583 AACATTATATCCCTTACAGGCTTTACTGTTTACAGATGGGAATGAATAATGTACAGCAA 1642
Db 1857 TCCATCATCTCTTACAAGGCTTACAGCAGACAGATGGCAATGAAGCAGTGCAGCAG 1916
QY 1643 CTCTTCAAGAGGTATCTCATTAGGTAATTCAGAGCCACACAGACAATTTGCTGGAAGCT 1702
Db 1917 ATTCTGAGTGAGAGTACACCTATACGTAATCTGATGTTGATTCAGACTCTTAGAGGCA 1976
QY 1703 GCAAGGCTGGAGATGCGAACTGTAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGC 1762
Db 1977 TCTAAGCTGGAGCTTGGAACTGTGAACCACTTTGCGACTCTCAAAATGTGAATGT 2036
QY 1763 AGAGACATTAAGGGGCTCAGTCTACACCCTTCAATTTTCAGAGCTGGGTATAACAGAGTG 1822
Db 2037 AGAGACTTAGAGGCGCGCATTTCCACGCCCTTACACTTCGCAAGGCTCAACCGCGTG 2096
QY 1823 TCCGTGTTGAATATCTGCTACAGCATGGAGCTGATGTGATGCTTAAGATAAAGGAGGC 1882
Db 2097 TCTGTTGATAGTACCTGCTACACCGCGGTGCGATGTTCCATGCGCAAGAGCAAGGGTGGC 2156
QY 1883 CTTGTACTCTTGTCAACATGATGTTCTTATGGACATTAATGAAGTTGCAAGACTTCTTGT 1942
Db 2157 TTGGTGGCCCTTCATATGCTGTTTCATATGCACTATGAGGTGGCTGAGCTTTTAGTA 2216

QY 1943 AAACATGGAGCAGTAGTTAATGTAGTGAATTTATGGAAATTTTACACCTTTTACATGAAGCA 2002
Db 2217 AGCATGGGGCTTCTGCAATGTGGCGACTTATGGAAATTTTACCCTCTCTCCATGAAGCA 2276
QY 2003 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGAGACCCCTACC 2062
Db 2277 GCAGCTAAAGGAAAATGAAATCTGCAAGCTCTTTTAAACATGGAGCAGATCCAAC 2336
QY 2063 AAAAAACAGGATGAAATACTCTCTTGGATCTTGTAAAGATGGAGATACAGATATT 2122
Db 2337 AAAAAACAGATGAAATACACCTTTGGATTTGTAAGGAAGGAGACACAGATATT 2396
QY 2123 CAAGATCTGTTAGGCGAGATGAGCTTTTGTAGTCTGCCAAGAGAGGTTGTTTAGCC 2182
Db 2397 CAGGACTTACGAAGGGGATGCTGCTTTGTTGATGCTGCCAAGAGGCTGCTCCCTGGCA 2456
QY 2183 AGAGTGAAGAGTTGCTTCTCTGATATAATGTAATTTGCGCGGATACCCAAAGGAGACAT 2242
Db 2457 AGAGTGCAGAGCTCTGTACCCAGAGAATATCAACTGCAGAGACACCCAGGGCAGAAAT 2516
QY 2243 TCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTCTTA 2302
Db 2517 TCAACCCCTCTGCACCTGGCAGCAGGCTATAATAAAGTGAAGTAGCTGAATATCTCTA 2576
QY 2303 CAACCGGAGCTGATGTGAATGCCCAAGCAAAAGGAGGACTTATTCCTTTTACATAATGCA 2362
Db 2577 GAGCATGGAGCTGATGTTAATGCCAGCAAGGTTGCTTTAATTCCTCTCTTCAATGCG 2636
QY 2363 GCATCTACGGGATGTAGATGTAGCAGCTCTACTATAAAGTATAATGATGTGTCAAT 2422
Db 2637 GCATCTATGGGATGTGATAGCAGCTTATTTGTAATAATACAAACAGCTGTGTAAAT 2696
QY 2423 GCCAGGACAAATGGGCTTTCAACACCTTTGCAGCAAGCAGCCCAAAAGGAGCAACACAG 2482
Db 2697 GCAACAGATAGTGGGCTTTACTCCCTCCATGAACGACGCCAGAGGAGGACGAG 2756
QY 2483 CTTTGTGCTTTGTTGTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAA 2542
Db 2757 CTGTGCGCCCTCTCTTAGCCCATGGTCAGAGCCCAACCATGAAGAACCCAGGAGGCCAG 2816
QY 2543 ACACCTTTAGATTTAGTTTTCAGCGGATGATCTCAGCGCTCTCTGACAGCAGCCATGCC 2602
Db 2817 ACAGCTCTGATCTGCAACAGCTGACGATATCAGAGCTTTGCTGATAGATGCCATGCC 2876
QY 2603 CCATCTGCTCTGCCCTCTTGTACAAAGCTCAAGTGTCTCAATGCTGTGAGAAGGCCAGGA 2662
Db 2877 CCAGAGGCCCTTACCTACCTGTTTTAAACCTCAGG-----CTACT 2915
QY 2663 GCCACTGCAGATGCTCTCTCTCAGGTTCCATCTAGCCCATCAAGCCTTTCTGCGAGCCAGC 2722
Db 2916 GTAGTGAGTGCCCTCTCTGATCTCACCAGCATCCACCCCTCTCGCTCTGGCTGCCAGC 2975
QY 2723 AGTCTTCACAACTTATCTGGGAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTTCAAGTGA 2782
Db 2976 AGCAATACAAACCTCACTGGCCCTTTAGCAGAGTTGGCGTAGGAGGAGCCTCCAAATGCA 3035
QY 2783 ACAGAGGCTGCTTCCAGTTTGGAGAAAAG-----GAGGTTCCAGGAGTAGATTTTAGC 2836
Db 3036 GGGATGCGCGCGGGAACACAGAAAGGAAGGAAGGAGAAAGTTGCTGTCTTGACATGAAT 3095
QY 2837 ATAACTCAATTCGTAAGGAATCTTTGGACTTTGAGCACCTAATGATATATTTTGAAGAGAA 2896
Db 3096 ATCAGCAATTTCTAAAAAGCCTTTGGCCTTGAACACCTTCGGGATATCTTTTGAACAGAA 3155
QY 2897 CAGATCACTTTTGGATGATTTAGTTGAGTGGGCGACAGAGCTGAAGGAGATTGGGAATC 2956
Db 3156 CAGATACATAGATGTTGGCTGATATGGTTCATGAAGAGTTGAAGAAATAGGCATC 3215
QY 2957 AATGCTTATGACATAGGCAACAACTAATTAAGAGGTGAGAGACTTATCTCGGACAA 3016
Db 3216 AATGCATATGGCAGCCCAACAAATTAATCAAGGAGTAGAAGAGCTCTTAGTGGACAA 3275

QY 3017 CAAGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAAACAATTTCTATAGAT 3076
Db 3276 CAAGCACCACCTCTATTTGACCTTTCACTGTGTTAATCAGGGAACGATTTTCTCGAT 3335
QY 3077 CTGCTCTGATGATAAGAGTTTCACTGTGTGGAGGAAGATGCAAAAGTACAGTTTCA 3136
Db 3336 CTGCTCCAGAAGATAAGATATCACTGAGTGGGAAGAGATGCAAAAGTACTATTTCA 3395
QY 3137 GAGCACAGATGAGGTCTATGAGGTGGAATCTTCAACAGATACAATATTTCTCAAGATT 3196
Db 3396 GAACACAGAGATGTTGATGCTGGCGGATCTTCAACAGATACAATGATGATTCGAATT 3455
QY 3197 CAGAAGTTTGTACAGAAGAACTATGGGAAGATACACTCACCGGAGAAAGATTTCT 3256
Db 3456 CAAAAGTTGTCAAGAAGAGTTGAGGAGCGGTCTCTGCCACGACAGAGAAGTGTCT 3515
QY 3257 GAAGAAACACACACCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3316
Db 3516 GAGGAGATCAACACCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3575
QY 3317 GCAATTATCCCAAGGCTTTGATGAAAGGATGCGTACATAGTGTGATGTTTGGAGCT 3376
Db 3576 GCCATTATCATAAAGGTTTGTAGCGGATGCGTACATAGGAGGATGTTTGGGCGC 3635
QY 3377 GGCATTTATTTGCTGAAACTCTTCAAAAGCAATCAATGATATGATGAATTTGGAGGA 3436
Db 3636 GGGATTTATTTGCTGAAACTCTTCAAAAGCAATCAATGATATGATGAATTTGGAGGA 3695
QY 3437 GGTACTGGGTGCTCCAGTTCAAGAGAGATCTGTTTACATTTGCCACAGGAGCTGCTC 3496
Db 3696 GGAACAGCTGCTTACACAGAGAGAGATGCTATATATGTCACAGACAAATGCTC 3755
QY 3497 TTTTGGCGGTAACTTTGGGAAAGTCTTTCTGCGATTCAGTGCATGAAATGGCACAT 3556
Db 3756 TTTCTGATAGTGACCTTTGGGAAATCTTTCTGCGATTTAGCACCATGAAATGGCCAC 3815
QY 3557 TCTCTCCAGGTATCACTCACTGCTAGTGTAGGCGGAGTGAATGGCTAGCATTAGCT 3616
Db 3816 GCGGCTCCAGGCGACCACTCAGTCTATGTTAGTACCGGCTCAATGGGCTGGCATATGCT 3875
QY 3617 GAATATGTTATTTACAGAGAGAGAGGCTTATCTCTGAGTATTTATTTACTTACCAGATT 3676
Db 3876 GAATATGCTATCTACAGAGAGAGAGGCTATCTCTGAGTATTTATTTACTTACCAGATT 3935
QY 3677 ATGAGGCGCTGAAG 3689
Db 3936 ATGAAGCCAGAAG 3948

RESULT 3

US-09-841-835-7
; Sequence 7, Application US/09841835
; Patent No. US200207695A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4491 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6..2027
; US-09-841-835-7

Query Match 38.2%; Score 1449.6; DB 10; Length 4491;
Best Local Similarity 65.3%; Pred. No. 0;
Matches 2397; Conservative 0; Mismatches 889; Indels 384; Gaps 3;
QY 383 TCCTCTGCTGCTTATCAATTAAGGTTTTGGGCGGAAAGAGTAGTTGAATATTTGCTTCAG 442
Db 657 TCTCCCTGCACCTTCGCTGCAGGTTTTGGAGGAAGAGATGTTGTAGAACACTTACTACAG 716
QY 443 AATGTGCAAGTGTCCAAACACGTCATGATGGGGCCCTTATTCCTCTTCATAATCATGCG 502
Db 717 ATGGGTGCTAATGTCCACGCTCGTCATGATGAGGTCTCATCCCGCTTCATAATGCTGT 776
QY 503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCTTTTGGCACATGTTGGAGAGCCCAATGCT 562
Db 777 TCTTTTGGCACATGCTGAGGTGTTGAGTCTGTTTATTTGCCAAGGAGCTGATCCAAATGCC 836
QY 563 CGAGATAATTTGAATTTATCTCTCTCCATGAAGCTGCAATTTAAAGGAAAGATGATGCT 622
Db 837 AGGATACTGAACATATACACCTCTGATGAGGTCTGATGAGGTCTGATGAGGTCTGATG 896
QY 623 TGCATTTGCTGTTTACAGCATGGAGCTGAGCCAAACCATCCGAAATACAGATGGAAGGACA 682
Db 897 TGCATTTGCTGCTGTCAGCAGCGAGCTGACCCAAACATTCGGAACACTGATGGGAATCA 956
QY 683 GCATTTGGATTTAGCAGATCCATCTGCCAAAGCAGTGTCTTACTGTTGAATATAAGAAAT 742
Db 957 GCGCTGGACCTGGCAGATCCCTTCAGCAAAAGCTGTCTTACAGGTGAATACAAAGAAC 1016
QY 743 GAACCTCTTAGAAAGTGCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTTACTCACACA 802
Db 1017 GAACCTCTTAGAAAGTGTCTGAGGAGTGGTAATGAAGAAAAAATGATGGCTTCTACTGCT 1076
QY 803 TTAATGTCAACTGCCACGCAAGTGTGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
Db 1077 CTAATGTGAATTTGCCATGCAAGTGTGGGCAAGTGTGCTCTTACATCTAGCAGCG 1136
QY 863 GGATATAACAGATTAAGATTTGTACAGTGTACTGCAACATGGAGCTGATGCTCATGCT 922
Db 1137 GGCTACAACAGATTCGAATAGTTCAGCTTCTTCTTACAGCATGGTGTCTGATGCTCA 1196
QY 923 AAAGATAAGGTGATCTGTTACCATTCACAAATGCTCTTCTTATGTTGTCATTATGAAGTA 982
Db 1197 AAAGACAAGGTGGACTTGTGCTCTTCTTCAATGCAATGCAATGCAATGCAATGCAATG 1256
QY 983 ACTGAATTTTGGTCAAGCATGGTGCCTTGTGTAATGCAATGCAATGCAATGCAATGCA 1042
Db 1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTCTTAAATGCAATGCAATGCAATGCAATG 1316

QY 1043 CTTCTTCATGAGGCGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAAAGTTAT 1102
DB 1317 CCACTGCACGAGGCTGCTTCCAAAGAACCGTGTAGAAGTCTGCTCTTGTACTTACGCCAT 1376
QY 1103 GTGTCAGACCCAACTGCTCAATTTGTCAAAATAAAGTGCTATAGACTTGGCTCCACAA 1162
DB 1377 GCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTCTGTGGATATGGCTCCAAC 1436
QY 1163 CCACAGTTAAAGAAATAGTAGCATATGAATTTAAAGGCCACTGCTGTGTCGAAGTGCA 1222
DB 1437 CCGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTACTTACAAAGCAGCC 1496
QY 1223 CGAAGAGCTGATCTTACTCGAATCAAAAACATCTCTCTCTGGAATGSGTAATTTCAAG 1282
DB 1497 AGAAGCAGACTAGTAAAGTTAAAAAACAACACTCGCTCTGGAATCATTAATTTCAAA 1556
QY 1283 CATCTCTAAACACATGAACAGACTTGCATTTGTGCTGCTCATCTCCATATCCCAAAAGA 1342
DB 1557 CAACCGAGTCTCATGAACAGCACTGCACCTGTGCTGTGGCTCTCTGCACTCCCAACGT 1616
QY 1343 AAGCAATATCTGAAGTTGCTTAAGAAAGAGCAACATCAATGAAGAGACTAAGAA 1402
DB 1617 AAACAAGTGACAGAAATTTGTTACTTAGAAAAGAGCAAAATTTAATGAAGAAATAAAGAT 1676
QY 1403 TTCTTGACTCTCTGCACGTGGCATCTGAGAAAGCTCATATATGATGTTGTGAAGTAGTG 1462
DB 1677 TTCACTACTCCCTGCTGATGTTGCAGCGAAGAGCCCATATATGATGTCATGGAAGTCTG 1736
QY 1463 GTGAACATGAAGCAAGGTTAATGCTCTGGATTAATCTTGGTCAGACTTCTCTACACAGA 1522
DB 1737 CATGAAGCATGGCGCAAGATGAATGCACTGGACACACCTTGGTCAGACTGCTTGGCATAGA 1796
QY 1523 GCTGCATATTTGGTGTCTATACAAACCTGCGCTACTCTGAGCTATGGTGTGATCCT 1582
DB 1797 GCGCCCTAGCAGGCCACCTGCAAGACCTGCGGCTCTGCTGAGTTACGGCTCTGACCCC 1856
QY 1583 AACATTATATCCCTTACAGGCTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAA 1642
DB 1857 TCCATCATCTCTTACAGGCTTACAGAGCTTACAGAGCAGACAGATGGCAATGAAGCAGTCCACAG 1916
QY 1643 CT----- 1644
DB 1917 ATTCTGAGTGTGAGTTAGGCTCTGACCCCTCCATCATCTCTTACAAAGGCTTACACAGA 1976
QY 1645 ----- 1644
DB 1977 GCACAGATGGCAATGAAGCAGTGCACAGATTTCTGAGTGGTCAATTCGTAGATAGTATC 2036
QY 1645 ----- 1644
DB 2037 ATTCTACTCAGCCTTAATGGTGTATTTAGACGGGAAGATTTAGAGAAATCTATCCA 2096
QY 1645 ----- 1644
DB 2097 GCATGCTTCACTGTCAACATGAAGAGTACACCTTATACGTACTTCTGTATGTTGATTATCG 2156
QY 1645 ----- 1644
DB 2157 ACTCTTAGAGGCATCTAAAGCTGGAGACTTGGAAACTGTGAAGCACTGTGAAGCACTTTCACACTCTCA 2216
QY 1645 -----C 1645
DB 2217 AATGTGAATTTAGAGACTTAGAGGCGCGCATTCACAGGCCCTTACACTTCGCAGCAGG 2276
QY 1646 CTCAAGAGGATCTCATTTAGTTAAATTCAGAGCAGCAGACAGAAATTTGCTGGAAGCTGCA 1705
DB 2277 CTACAACAGAGTACACCTATACGTACTTCTGATGTTGATTATCGACTCTTAGAGGCACTCT 2336
QY 1706 AAGCTCGAGATGTCGAACCTGTAATAAACTGTACTGTTACAGAGTGTCAACTGCGAGA 1765
DB 2337 AAAGCTGGAGACTTGGAAACTGTGAAGCACTTTGCAGCTCTCAAAATGTTGAATTTGAGA 2396

QY 1766 GACATTGAAGGGCGTCAGTCTACACCACTTCAATTTGCAGCTGGGTATATAACAGAGTGTCC 1825
DB 2397 GACTTAGAGGGCGGCATTCACGCGCTTACACTTCCAGCAGGCTACAACCGCGTCT 2456
QY 1826 GTGGTGAATATCTGTACAGCATGGAGCTGATGCTGAAGTAAAGATAAGAGAGGCTT 1885
DB 2457 GTTGTAGAGTACCTGTACACCGGTCGCGATGTCATCCAAAGACAAAGGTTGGCTG 2516
QY 1886 GTACCTTTGCACAAATCATGCTTTATGGACAATTATGAAGTTTGCAGAACTTCTTGTTTAAA 1945
DB 2517 GTCCCTTCAATAGCTGCTTCAATATGGACATATGAGTGCTGAGCTTTTAGTAAGG 2576
QY 1946 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAATTTTACACCTTTACATGAAGCAGCA 2005
DB 2577 CATGGGCTTCTGTCATATGTCGCGGACTTATGGAATTTTACCCCTCTCCATGAAGCAGCA 2636
QY 2006 GCAAAAGGAAAATATGAATTTTGCAAACTTCTGCTCCAGCATGGTCGAGACCCCTACCAAA 2065
DB 2637 GCTTAAAGGAAAGTATGAATCTCGAAGCTCTTTTAAACATGGAGCAGATCCAACATAA 2696
QY 2066 AAAAAAGGAGTGAATACTCTCTTTGGATCTTGTAAAGATGGAGATACAGATATTCAA 2125
DB 2697 AAGAACAGAGATGAATAACACTTTGGATTTGTAAGGAAGGAGACACAGATATTTCAG 2756
QY 2126 GATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTTAGCCAGA 2185
DB 2757 GACTTACTGAAGGGGATGCTGCTTTGTTGGATGCTGCCAAGAGGGCTGCTGGCAAGA 2816
QY 2186 GTCAAGAGTTGCTTCTCTCTGATATGTAATTTGCCGCGTATACCAAGCAGACATCA 2245
DB 2817 GTGCAGAGTCTGTACCCAGAGAAATATCACTGTCAGAGACACCCAGGAGCAAAATCA 2876
QY 2246 ACACCTTTACATTTAGCAGCTGTTATAATTTTAAAGTTGCAGAGTATTTTGTACAA 2305
DB 2877 ACCCTCTGCACCTGGCAGCAGCTATATAACCTGGAAGTAGCTGAATATCTTCTAGAG 2936
QY 2306 CACGGAGCTGATGTAATGCCAAGACAAAGAGGAGCTTATTCCTTTACATAATGACAGA 2365
DB 2937 CATGGAGCTGATGTAATGCCAGCAAGGGTGTAAATTTCTCTTCAATAATGCGGCA 2996
QY 2366 TCTTACGGGATGTAGATGTAGCAGCTCTACTAATAAGTATAATGATGTGTCAATGCC 2425
DB 2997 TCTTATGGGATGTTGACATAGCGGCTTTATTGATATAATACACAGCTGTGTAATATGCA 3056
QY 2426 ACGACAAATGGGCTTTTACACCTTTTGCAGAGCAGCCCAAAAGGAGGACACAGACTT 2485
DB 3057 ACAGATAAGTGGGCTTTTACTCCCTCCATGAAGCAGCCCAAAAGGAGGACGAGCTG 3116
QY 2486 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAGGACAAACA 2545
DB 3117 TGGCCCTTCTTCCCTAGCGCATGTGTCAGACCCCACTGTAAGAACCCAGGAGGCGCAGC 3176
QY 2546 CTTTAGATTTAGTTTTCAGGGGATGATGTCAGCGCTCTTCTGACAGCAGCCTGCCCCCA 2605
DB 3177 CCTGTGATCTGGCAACAGCTGAGATATCAGAGTTTGTCTGATAGATGCTATGCCCCCA 3236
QY 2606 TCTGCTCTGCCCTCTGTTTACAAGCCTCAAGTGTCTCAATGGTGTGTAAGGCCAGGAGCC 2665
DB 3237 GAGGCCCTTACTACTCTTTTAAACCTCAGG-----CTACTGTA 3275
QY 2666 ACTGCAGATGCTCTCTTCTAGGTCATCTAGCCCATCAAGCCCTTCTGACGCCAGCAGT 2725
DB 3276 GTGAGTGGCTCTGTGATCTCACCAGCATCCACCCCTCTCGGCTGCGCCAGCAGC 3335
QY 2726 CTTGACAACTTATCTGGGAGTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAAACA 2785
DB 3336 ATAGACAACCTCACTGGCCCTTTAGCAGAGTTGGCCGTAGGAGGAGCCTCCAATGCGAGG 3395
QY 2786 GAGGCTCTTCCAGTTTGGAGAAAAAGAG-----GTTCCAGGAGTAGATTTTAGCAATA 2839
DB 3396 GATGGCGCGCGGAACAGAAAGGAGGAGGAGGAGTGTGTGTCTTGTACATGAATATC 3455
QY 2840 ACTCAATTCGTAAAGAACTCTTGGACTTGCAGCACCTAATGGATATATTTGAGAGAGAACAG 2899

Db 3456 AGCCAAATTTCTAAAGAGCTTGGCTTGAACACCTTCGGGATATCTTTGAAACAGACAG 3515
Qy 2900 ATCACTTTGGATGATTTAGTTAGATGGGCGACAGAGCTGAAGAGATTTGGAATCAAT 2959
Db 3516 ATTACACTAGATGTTGGCTGATAGGCTCATGAAGAGTTGAAAGAAATAGGCATCAAT 3575
Qy 2960 GCTTATGACATAGGACACAACTAAATTAAGGAGTTCGAGAGACTTATCTCCGGACACAA 3019
Db 3576 GCATATGGCCGCCCAAAATTAATCAAGAGATAGAGACTCTTAGTGGGACACAA 3635
Qy 3020 GGTCTTAACCCCAATTTAATTTGAACACCTCTCTGTAGTGAACAAATTTCTTATAGATCTG 3079
Db 3636 GGCACCAATCTTATTTGACTTTTCACTGTCTTAATCAGGGAACGATTTCTGTGATCTT 3695
Qy 3080 TCTCTGATGATAAAGAGTTTCACTCTCTGGAGGAGATGCAANGTACAGTTTCGAGAG 3139
Db 3696 GCTCCAGAGATTAAGATATACGTACGTGGAAGAGAGATGCAAGTACTATTTCGAGAA 3755
Qy 3140 CACAGAGATGAGGTCATGCGAGGTGGAATCTTCAACAGATACAAATTTCTCAAGATTCAAG 3199
Db 3756 CACAGAGATGGTGAATGCTGGCGCATCTTCAACAGATACAAATTTCTCAAGATTCAAG 3815
Qy 3200 AAGTTTGTAAACAGAACTATGGGAAGATACACTTCCCGGAGAAAGAAAGTTTCTGAA 3259
Db 3816 AAGTTTGTAAACAGAACTATGGGAGCGGTCTGCCACCGACAGAGAAAGTCTCTGAG 3875
Qy 3260 GAAACACCAACCATGCAATGAAGATGCTATTTTCATGGGTCTCTTTTGTGAATGCA 3319
Db 3876 GAGAATCACACCATGCAATGAGGCGCATGTTTTCATGGTCTCTTTTCAATTAATGCC 3935
Qy 3320 ATTATCCCAAGAGCTTTGATGAAAGGATGCGTACATAGGTGATGTTTGGAGCTGGC 3379
Db 3936 ATTATTCATAAGGTTTGTAGTGGCGCATGCATACATAGGAGGAATGTTTGGGCGCGG 3995
Qy 3380 ATTATTTTGTGAAACTCTTTCCAAAAGCAATCAATATGATATGAAATTTGGAGGAGGT 3439
Db 3996 ATTATTTTGTGAAACTCTTTCCAAAAGCAATCAATATGATATGAAATTTGGAGGAGGA 4055
Qy 3440 ACTGGTGCAGGTCACAAAGACAGATCTGTTTACATTTGCCACAGGAGCTGCTCTTT 3499
Db 4056 ACAGGCTGCCCTTACACAAAGGAGGTCATGCTATATGTCACAGACAAATGCTCTTC 4115
Qy 3500 TGCCGGGTAACTTTGGGAAGTCTTTTCTGCAAGTTTCAAGTGAATGAAATGGCACATTTCT 3559
Db 4116 TGTAGAGTGACCTTTGGGAATCTTTTCTGCAAGTTTCAAGTGAATGAAATGGCACAGCG 4175
Qy 3560 CTTCCAGGTCATCACTCACTGCTAGTGGTGGGCGGCTGTAATGGCTTACGATTTAGCTGAA 3619
Db 4176 CTTCCAGGTCATCACTCACTGCTAGTGGTGGGCGGCTGTAATGGCTTACGATTTAGCTGAA 4235
Qy 3620 TATGTTATTTACAGAGGACACAGCTTATCTGAGTATTTAATTTACTTACCAGATTTATG 3679
Db 4236 TATGTTATTTACAGAGGACACAGCTTATCTGAGTATTTAATTTACTTACCAGATTTATG 3679
Qy 3680 AGGCTTGAAG 3689
Db 4296 AAGCCAGAAG 4305

RESULT 4

US-09-841-835-9
; Sequence 9, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor

; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4657 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6...2855

US-09-841-835-9

Query Match 34.2%; Score 1297.8; DB 10; Length 4657;

Best Local Similarity 62.9%; Pred. No. 0;

Matches 2400; Conservative 0; Mismatches 907; Indels 508; Gaps 5;

Qy 383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGGGAAACACGCTAGTTGAATATTTGCTTCAG 442
Db 657 TCTCCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGATGTTGTAGAACACTTACTACAG 716
Qy 443 AATGGTCAAGTGTCCAAAGCACGATGATGGGGGCTTATCTCTTTCATATGCAATGC 502
Db 717 ATGGGTCTAATGTCACGCTCGTGTATGATGGAGTCTCATCCGCTTCATATGCTGT 776
Qy 503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCTTTTGGACATGCTGCAGACCCCAATGCT 562
Db 777 TCTTTTGGTCATGCTGAAGTAGTCAATCTCTTTTGGACATGCTGCAGACCCCAATGCT 836
Qy 563 CGAGATTAATGGAATTTACTCTCTCCATGCAAGCTGCAATTAAGGAAGATTTGATGTT 622
Db 837 AGGATAACTGGAATTTACTCTCTCCATGCAAGCTGCAATTAAGGAAGATTTGATGTT 896
Qy 623 TGCATTTGCTGTTTACAGCATGGAGCTGAGCCAAACCATCCGAAATACAGATGGAAGACA 682
Db 897 TGCATTTGCTGCTGACAGCAGGAGCTGACCCAAACCATTCGGAACACTGATGGAAATCA 956
Qy 683 GCATTTGATTTAGCAGATGCCATCTGCCAAAGCAGTGTCTTACTGGTGAATATAAGAAAGAT 742
Db 957 GCGCTGGACCTGGCAGATCTCTCAGCAAAAGCTCTCTTACAGGTGAATAAGAAAGAC 1016
Qy 743 GAATCTTTAGAAAGTCCAGGAGTGGCAATGAAGAAAAATGATGGCTCTACTACACCA 802
Db 1017 GAATCTTTAGAAAGTCCAGGAGTGGCAATGAAGAAAAATGATGGCTCTACTACACCT 1076
Qy 803 TTAATGTCAACTGCGAGCAAGTGTGGCAGAAAGTCAACTCCATTACATTTTGGCAGCA 862
Db 1077 TTAATGTCAACTGCGAGCAAGTGTGGCAGAAAGTCAACTCCATTACATTTTGGCAGCG 1136

Qy	863	GGATATTAACAGAGAAAGAATTGTACAGCTGGTTACTTGCACAATGGAGCTGTATGTCCTCATGCT	922
Db	1137	GGCTACACAGAGTGGATGGCTTTCATATGATGCTTCTTCTTTCAGCATGGTGCTGATGTTTCATGCA	1196
Qy	923	AAAGATAAAGGTGATCTGGTACCAATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA	982
Db	1197	AAAGCAAAAGTGGACTTGGCGCTTTCATAATGATGTTTCATATGGACATATTAATGAAGTC	1256
Qy	983	ACTGAACCTTTTGGTCAAGCATGGTGCTGTGTAATGCAATGGACTTGTGGCAATTCACAT	1042
Db	1257	ACAGAACTGCTACTAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGCTTTACT	1316
Qy	1043	CCTCTTCATGAGGAGCTTCTTAAGAACAGGGTTGAAGTATGTTCTTCTTCTCTTAAGTTAT	1102
Db	1317	CCATGCAAGGCTGCTTCCAAAGAACCGTGTAGAAGTCTGCTCTTGTGTTACTTTAGCCAT	1376
Qy	1103	GGTSCAGACCCACACCTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCACACA	1162
Db	1377	GGCGCTCATCCTACGTTAGTCAACTGCCATGCCAAGAAAGTGTGTGATATGGCTCCAACAT	1436
Qy	1163	CCACAGTTAAAGAAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGTGCA	1222
Db	1437	CCGGAGCTTAGGGAGAGATGTACTATGATTTAAAGTCAATCTTTACTACAAGCAGCC	1496
Qy	1223	CGAAGACTGATGTTACTCGAATCAAAAACATCTCTCTCGAAATGGTGAATTTCAAG	1282
Db	1497	AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGAAATCAATTAATTTCAA	1556
Qy	1283	CATCCTCAACACATGAACACGATTGCAATGTGTGCTGTCATCTCCATGCAATGCCAAAGA	1342
Db	1557	CAACCGCAGTCTCATGAACACGACTGCACTGTGCTGTGGCTCTCTGCAATGCCAACGCT	1616
Qy	1343	AAGCAAAATATGTCAACTGTTGCTTAAGAAAAGGAGCAACATCAATGAAAGACATTAAGAA	1402
Db	1617	AACNAGTGACAGAAATGTTACTTAGAAAAGGAGCAAAATGTAATGAAAATAAAGAT	1676
Qy	1403	TCTTGACTCCTCTGCACGTGGCATCTCAGAAAGCTCATATGATGTTGTGAAGTAGTG	1462
Db	1677	TTCATGACTCCCCCTGCATCTGCAGCCGAAAGAGCCCATATGATGTCATGGAAGTCTG	1736
Qy	1463	GTGAACATCAAGCAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA	1522
Db	1737	CATAGCATGGCGCCAAAGATGAATGCATGGACACCCCTGGTCAGACTGTTGTCATAGA	1796
Qy	1523	GCTGCATATTTGGTCATCTACAAACCTCGCGCCTACTCCTGAGCTATGGTGATGCTCT	1582
Db	1797	GCGGCCCTAGCAGGCCACCTGCGAGACCTTGGCGCCTCTCTGTAGTTAGGCTCTGACCC	1856
Qy	1583	ACATATATCCCTTCAGGCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAA	1642
Db	1857	TCCATCATCTCCTTACAAGGCTTTCACAGCACAGATGGCAATGAAGCAGTGCAGCAG	1916
Qy	1643	CTCCTCCAGAGGGTATCTCATTAGGTAAATTCAGAGGACAGACANITGCTTGAAGCT	1702
Db	1917	ATTCTGAGTGAGAGTACACCTATACGTACTTCTGATCTTGATTTACGACTCTTTAGAGCA	1976
Qy	1703	GCAAAAGCTGGAGATGTCGAAACTTAAAAAACTGTGTACTGTCAGAGTGTCACTG	1762
Db	1977	TCTAAAGCTGGAGACTTGGAACTGTGAAGCACTTTGCAGCTCTCAAAATGGAATTTG	2036
Qy	1763	AGAGACATTTGAAGGGCGTCACTGTACACCACTTTCATTTGCACTGGGTATACAGAGTG	1822
Db	2037	AGAGACTTAGAGGGCGGCAATTCACAGGCCCTTACACTTCGACAGGCTACACCCGCTG	2096
Qy	1823	TCCGTGGTGGATATCTGCTACAGCATGGAGCTGATGTGCATGCTTAAGATAAAGGAGC	1882
Db	2097	TCTGTGTTAGAGTACTGCTACACCAAGGTGCCATGCTCAIGCCAAAGACAAAGGTGCG	2156
Qy	1883	CTTGTAACCTTTGCACAAATGCAATGTTCTTATGGACATTAATCAAGTTGCAGAACTTCTGT	1942
Db	2157	TTGGTGCCCCCTCAATATGCCCTGTTTCATATGGACACTTTTGGGTGGCTGAGCTTTAGTA	2216
Qy	1943	AAACATGGAGCAGTAGTTAATGTAAGCTGATTTATGAAATTTACAGCTTTACATGAAGCA	2002

Db	2217	AGGCATGGGCTTCTGTCAAATGTGGCGGACTTATGMAATTTACCCCTCTCCATGAAGCA	2276
Qy	2003	GCACCAAAAGGAANAATAGAAATTTGCAAACTTGTCTCCAGCATGGTGCAGACCCCTACC	2062
Db	2277	GCAGCTAAAGGGAAGTATGAAATCTGCAAGCTCCTTTTAAAAATGAGAGATCCCAACT	2336
Qy	2063	AAAAAACAAGGATGAAATACCTCTTTGGATCTTGTTAAAGATGGAGATACAGATATT	2122
Db	2337	AAAAGAACAAGATGGAATAACCTTTGGATTTGGTAAAGGAAGGAGACACAGATATT	2396
Qy	2123	CAAGATCTGCTTAGGGGAGATGCAGCTTTTGCTAGATGCTGCCAAGAAGGTTGTTTAGCC	2182
Db	2397	CAGGACTTACTGAAGGGGATGCTGCTTTGTTGGATGCTGCCAAGAAGGCTGCCTGGCA	2456
Qy	2183	AGAGTGAAGAAGTTGTCTTCTCTGATAATGTAAATTTGCCCGATACCCAAAGGCAGACAT	2242
Db	2457	AGAGTGCAGAAGCTCTGTACCCCGAGAGAATATCAACTGCAGAGACACCCAGGCGAGAAT	2516
Qy	2243	TCACACCTTTACATTTAGCAGCTGGTTATATATTTAGAACTTGCAGAGATTTTGTTA	2302
Db	2517	TCAACCCCTCTGCACCTGGCGAGGCTATATATACCTTGGAAAGTAGCTGAATATCTTCTA	2576
Qy	2303	CAACACGAGCTCATGTGAATGCCAAGACAAGCAGGACTTATTTCCTTTACATAATGCA	2362
Db	2577	GAGCATGGAGCTCATGTTAATGCCCAGGACAGGCTGGTTAATTCCTCTTCATATATGCG	2636
Qy	2363	GCATCTTTACGG-----	2373
Db	2637	GCATCTTATGGGGCTGCCTGGCAAGAGTGCAGAAGCTCTGTACCCCGAGAGAATATCAAC	2696
Qy	2374	-----	2373
Db	2697	TGCAGAGACACCCAGGGCGAAATTCACCCCTCTGCACCTGGCAGCAGGCTATAATAAC	2756
Qy	2374	-----	2373
Db	2757	CTGGNACTAGCTGAATATCTTCTAGAGCATGGAGCTGATGTTAATGCCAGGACAGGGT	2816
Qy	2374	-----	2373
Db	2817	GGTTTAATTCCTCTTCATATAGCGCATCTTATGGTAGTAAAAAGTTGGATTCCAGACC	2876
Qy	2374	-----	2373
Db	2877	TCCTTCCAGCTTGTGTTAATGATTAATAGACCATGCATGTGGAATTTGCATTAACTA	2936
Qy	2374	-----GCATGTAGATGTAGCAGCTCTACTATAAAGTATAA	2409
Db	2937	ATGTAAGGCATTATAAAAATGCAAGCATGTTGCATACGGGCTTTTATGTATAAATACAA	2996
Qy	2410	TGCATGTGTCATGCCACGGACAAATGGGCTTTACACCTTTGCACGAAGCAGCCCAAAA	2469
Db	2997	CACGTGTGTAATGCAACAGATAAGTGGGCGTTTACTCCCGCTCCATGAACGCCCGAGAA	3056
Qy	2470	GGACACCAACACAGCTTTGTGCTTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAA	2529
Db	3057	AGAAAGGACGCGAGCTGTGGCGCCCTCCTCTAGCGCATGGTGCAGACCCACCATTGAAGA	3116
Qy	2530	TCAGGAAGGACAAACACCTTTTAGTTTATGTTTCAGCGGATGATGTCAGCGCTCTTC-	2585
Db	3117	CAAGNAGCCAGACGCCCTGCGGATCTGGCAACAGCTGCAGATATCAGAGCTTTCATGT	3176
Qy	2586	-----	2585
Db	3177	TGACATAGCGGCTTTATTGATAAAAATACAACACGCTGTGTAATGCAACAGATAAGTGGCG	3236
Qy	2586	-----TGACAGAGCGCATGCCCCCATCTGCT	2611
Db	3237	GTTTACTCCCTCCATGAAGCAGGCCAGAAAAGGAAGCAGCGCTGTGTGCCCTCTCTCT	3296
Qy	2612	CTGCCCTCTTGTTACAAAGCTCAAG-----	2636

Db 3297 AGCGCATGTGCGAGACCCACCACCATGAAGAACAGGAGCCAGACGCGCTCTGGATCTGGC 3356
Qy 2637 -----TGCCTCAATGGTGTGAGAACCCAGGAGCCACTGC--- 2670
Db 3357 AACAGCTGAGATATCAGAGCTTTGCTGATAGATCCATGCCCCAGAGGCGCTTACCCTAC 3416
Qy 2671 -----AGATGCTCTCTCTTCCAGGTCCATCTAGGCC 2700
Db 3417 CTGTTTTAAACCTCAGGCTACTGTAGTAGTGCCTCTCTGATCTCACCAGCATCCACCCC 3476
Qy 2701 ATCAAGCCCTTCGAGCCAGCAGCTCTTGACAACCTATCTGGAGTGTTCGAAACTGTC 2760
Db 3477 CCTCTGCTCTGCGCTGCGCAGCAGCATAGACAACCTCACTGGCCCTTTAGCAGAGTTGC 3536
Qy 2761 TTCAGTAGTTAGTTCAGTGGACAGAGGTGCTTCCAGTCTTGGCAAAAAG-----CA 2814
Db 3537 CGTAGGAGAGCCCTCAATGCGAGGGATGGCGCGGGAACAGAAAGGAAGGAGA 3596
Qy 2815 GGTTCAGGAGTAGATTTTAGCATACTCAATTCGTAAAGAACTTTGGACTTGGACACCT 2874
Db 3597 AGTTGCTGCTTGGACATGAATATCAGCCAATTTCTTAAAAAGCCCTTGGCCCTTGAACACCT 3656
Qy 2875 AATGGATATATTGAGAGAGACAGATCACTTTGGATGTATTAGTTGAGATGGGGACAA 2934
Db 3657 TCGGGATATCTTTGAACACAGACAGATTACACTAGATGTGTTGGCTGATATGGTGCATGA 3716
Qy 2935 GGAGCTGAAGGACATTTGAATCAATGCTTATGGACATAGCACAACCTAAATTAAGGAGT 2994
Db 3717 AGAGTTGAAGAATAAGGATCAATGCAATGGCCACCCCAAAATTAATCAAGGAGT 3776
Qy 2995 CGAGAGACTTATCTCGGACCAAGAGTCTTAACCCATATTTAACTTTGAACACCTCTCG 3054
Db 3777 AGAAGACTCTTAGTGGACAAAGCACCACCAATCTTATTTCACTTTTCACTGTCTTAA 3836
Qy 3055 TAGTGGAACTTCTTATAGATCTGCTCTGATGATATAAGAGTTTCAGTCTGTGGAGGA 3114
Db 3837 TCAGGGAACGATTTTGTCTGATCTTCTCCAGAGATAAAGAAATATCAGTCAGTGAAGA 3896
Qy 3115 AGAGATGCAAGTACAGTTTCGAGACACAGATGAGATGAGCTCATGAGGTGGAATCTTCAA 3174
Db 3897 AGAGATGCAAGTACATTCGAGAACACAGATGATGGTGAATGCTGGCGGATCTTCAA 3956
Qy 3175 CAGATCAATATTTCTCAAGATTCAGAGGTTTGTAAACAAAGAACTATGGGAAAGATACAC 3234
Db 3957 CAGATCAATATTCGAAATTCAAAAGTTGTCACAAAGAGTTGAGGGAGCGGTTCTG 4016
Qy 3235 TCACCGGAGAAAAGAGTTTCTCAAGAAACACACACCATGCAATGCAAGATGCTATT 3294
Db 4017 CCACCGACAGAGGAAGTGTCTGAGGAGAAATCAACACCATCAATGAGCGCATGTTGT 4076
Qy 3295 TCATGGTCTCCTTTTGTGAATGCAATTTATCCAAAGGCTTTGATGAAAGGATGCGTA 3354
Db 4077 TCATGGTCTCCTTTTCAATGCAATTTATCAAGGGTTTGTAGCGGATGCGAT 4136
Qy 3355 CATAGTGTGATGTTTGGAGCTGGCATTTATTTGCTGAAAACCTCTTCCAAAGCAATCA 3414
Db 4137 CATAGGAGGAATCTTTGGGCGCGGATTTATTTTGTGAAAACCTCTTCAAAAAGCAACCA 4196
Qy 3415 ATATGATATGGAATTTGGAGGAGTACTGGGTCTCCAGTTTCACAAAGACAGATCTTGTTA 3474
Db 4197 ATATGTTTATGGAATTTGGAGGAGGAACAGGCTGCCCTTACACAAAGGAGGTCTATGCTA 4256
Qy 3475 CATTTGGCCACAGCAGCTGCTCTTTTGGCGGTTAACTTTGGGAAAGCTTTTCTGCGAGTT 3534
Db 4257 TATATGTCACAGACAATGCTCTTCTGTAGATGACCCCTTGGGAAATCTTTCTGCGAGTT 4316
Qy 3535 CAGTGAATGAAATGGCAATCTCTCTCCAGGTCTATCACTCAGTCACTGGTAGGCCAG 3594
Db 4317 TAGCACCATGAAATGGCCACGCGCTCCAGGCGACCTCAGTCACTTGGTAGACCGAG 4376
Qy 3595 TGTAAATGGCTTAGCATAGCTGAATATGTTATTTACAGAGGAGAACAGCTTATCCTGA 3654
Db 4377 CGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCAATACCCAGA 4436

Qy 3655 GTATTTTATTTACCATGATTATGAGCCCTGAAG 3689
Db 4437 GTATCTTATCTACTTACCAGATCAITGAAGCCAGAAG 4471

RESULT 5

US-09-964-899-40
; Sequence 40, Application US/09964899
; Patent No. US2002017446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-964-899-40

Query Match 26.2%; Score 994; DB 9; Length 2409;
Best Local Similarity 65.4%; Pred. No. 1.8e-254;
Matches 2174; Conservative 0; Mismatches 20; Indels 1128; Gaps 4;

Qy 403 AGTTTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAATGGTCCAAGTGTCCAAAGC 462
Db 198 AGGTTTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAATGGTCCAAGTGTCCAAAGC 257
Qy 463 AGCTGATCATGGGGCCCTTATTCCTCTTCAATCAATGCTCTTTTGGTCTATGCTGAAGT 522
Db 258 AGCTGATCATGGGGCCCTTATTCCTCTTCAATCAATGCTCTTTTGGTCTATGCTGAAGT 317
Qy 523 AGTCAATCTCCTTTTGGACATGTTGACAGCCCAATGCTCGAGATAATTGGAATATATAC 582
Db 318 AGTCAATCTCCTTTTGGACATGTTGACAGCCCAATGCTCGAGATAATTGGAATATATAC 377
Qy 583 TCCTCTCCATGAAGTGCATTAAGGAAAGATTTGATGCTTTGCTGCTTACAGCA 642
Db 378 TCCTCTCCATGAAGTGCATTAAGGAAAGATTTGATGCTTTGCTGCTTACAGCA 437
Qy 643 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTAGCAGATCC 702
Db 438 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTAGCAGATCC 497
Qy 703 ATCTGCCAAAGCAGTGTCTTACTGTTGAATAT-----AAGAAAGATGA 744
Db 498 ATCTGCCAAAGCAGTGTCTTACTGTTGAATATTTGCTTATTCAGGAAAGCCCTGT 557
Qy 745 ACTCTTAGAAAGTGCAGAGTGGCAATCAAGAAAAATGATGCTCTACTCACACCAATT 804
Db 558 AAAGAACCAACCTTGGCAGAGTGGCAATGAAGAAAAATGATGCTCTACTCACACCAATT 617
Qy 805 AAATGCTCAACTGCCACGCAAGTGAATGGCAGCAAGTCAACTCCATTTACATTTTGGCAGCAGG 864
Db 618 AAATGCTCAACTGCCACGCAAGTGAATGGCAGCA----- 648
Qy 865 ATATAACAGAGTAAAGATTTGATACAGCTGTTTACTGCAACATGGAGCTGATGCTCATGCTAA 924
Db 649 ----- 648
Qy 925 AGATAAGGTGATCTGGTACCATTACACAATGCCCTGTTTATGCTCATTTATGAGTAAC 984
Db 649 ----- 648

US-09-964-899-42
; Sequence 42, Application US/09964899
; Patent No. US2002017446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 5175
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-964-899-42

Query Match 2.1%; Score 80.8; DB 9; Length 5175;
Best Local Similarity 45.5%; Pred. No. 5.4e-11;
Matches 390; Conservative 0; Mismatches 447; Indels 21; Gaps 2;

QY 1356 AACTGTTCTGAAGAAAGAGCAACATCAATGAAAGACTAAAGAAATTCCTGACTCCTC 1415
Db 740 AACTATTCTCGATCGAGGAGTAAATCGATGCCAAACACAGGAGGTGCTGACACAC 799

QY 1416 TGCAGTGGCATCTGAGAAGCTCATATGATGTGTTGAAGTAGTGTGGAACATGAAG 1475
Db 800 TGCAGTGGAGCAAGGAGTGGCCACGAGCAGGTGGTGAATGTTGTTGATCGAGCTG 859

QY 1476 CAAGGTTATGCTCTGATAATCTTGCTCAGACTCTCTACAGAGCTGCATATTCTG 1535
Db 860 CCCCATCTTTCAAAACCAAGATGATATTCTCCATTGCATGCCACACAAAGGG 919

QY 1536 GTCACTACAAACCTGGCCCTACTCCTGAGCTATGGGTGTGATCTTAAACATTATATCC 1595
Db 920 ATCATTTAAAGCTGGCTCCAGCTTCTCCAGCATATGTACCCGTGGATGATGTCACCA 979

QY 1596 TTCAGGGCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCTCCAAAGG 1655
Db 980 ATGACTACCTGACCTGCTTACAGTGGCTGCCACTGTGGCCATTACAAAGTTCGCAAG 1039

QY 1656 GTATCTCATTAAGTAATTCAGAGCAGACACAAATTCCTGGAAGCTGCAAGGCTGGAG 1715
Db 1040 TTCTCTTGGAATAAGAAAGCTAACCCCAATGCCAAGCCCTGAATGGCTTTACCCCTCTC 1099

QY 1716 ATGTCGAA-----ACTGTAAAAAAGCTGTGACTGTTCAGAGTGTCA 1757
Db 1100 ATATTGCTCTCAAGAAAGATCGAATTAAGTAATGGAATCTCTTGAAGACACGGTGCAT 1159

QY 1758 ACTCAGAGACATGAAGGGCTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1817
Db 1160 CCATCCAGCTGTAAACGAGAGAGAAACAGACACTGCACATGGCAGCTCGCTCCGGCC 1219

QY 1818 GAGTGTCCGGTGGATATCTGCTACAGCATGAGCTGATGTGCATGCTTAAAGATAAAG 1877
Db 1220 AAGCTGAAGTGTGCGGTATCTGTGTACAAAGCAGGAGCTCAGGTAGAAGCTAAAGCTAAG 1279

QY 1878 GAGGCCCTGTACCTTTCACAAATCATGTTCTTATGACATATTAAGTGTGCAAACTTC 1937
Db 1280 ATGACCAACACCACTCCCAATTCAGCCGCACTGGGGAAAGCAGACATAGTACCAAGC 1339

QY 1938 TTCTTAACATGGAGCAGTAGTAAATAGTGTATTTATGAAATTTACACCTTTACATG 1957
Db 1340 TGTTCAGCAAGGGCATCTCCAAATGACGACCACTTCTGGGTACACCCCACTTCACC 1399

QY 1998 AAGCAGCAGCAAAAGGAAATATGAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2057
Db 1400 TTTCCCGCCGAGGGGAGGAGGATGGCGCGTCTCTTTGTTGATCATGGAGCGTCTT 1459

QY 2058 CTACCAAAAAAACAGGGATGGAATACTCCTTGGATCTTGTAAAGATGGAGATACAG 2117
Db 1460 TATCTATATAACAACAAGAGCGGGCTAACACCACTGCATGTAGCTGCACATTTACGATAATC 1519

QY 2118 ATA---TTCAAGATCTGTTAGGGAGATGACGCTTTTGTAGATGCTGCCAAGAGCGTT 2174
Db 1520 AGAAAGTGCCTTCTGCTTTTGGACCAAGGAGCCTCAGCTCAGCGAGCGCCAAAGAATG 1579

QY 2175 GTTTAGCCAGAGTGGAAGA 2192
Db 1580 GTTATACGCCACTGCACA 1597

RESULT 13
US-09-908-711-12
; Sequence 12, Application US/09908711
; Patent No. US20020045230A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P128
; CURRENT APPLICATION NUMBER: US/09/908,711
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US01/01360
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,867
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01344
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01345
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,888
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01329
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,905
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01354
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,891
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01339
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01340
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,874
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01334
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,898
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01320
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,853
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01349
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,902
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01239
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,870
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01348
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,882
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01347

```

; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,896
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01307
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,864
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01341
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,856
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01336
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,868
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01312
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/209,457
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (13)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1287)
; OTHER INFORMATION: n equals a,t,g, or c
; US-908-711-12

```

	Query Match	2.0%;	Score 77;	DB 10;	Length 1299;	
	Best Local Similarity	45.5%;	Pred. No.	2.5e-10;		
	Matches 337;	Conservative 4;	Mismatches 394;	Indels 6;	Gaps 2;	
Qy	1365	TAGAAAAAGGACAAACATCAATGAAGAAGCTATAAAGAAATTCTTGACTCCCTCGTCACGGTGG	1424			
Db	203	TAATAAAATGGAGTTGCATCAACAATTGCCAATCAGAATGGGTGGAACGGCTCTCCACCCTTG	262			
Qy	1425	CATCTCAGAAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACACATGAAGCAAAGCTTA	1484			
Db	263	CTTCCAAGAAAGGCCATGTAGAGTTGTTTTCTGAGCTGCTGCAGAGAACGCCAATGTGG	322			
Qy	1485	ATGCTCTGGNTRATCTTGCTGACACTTCTCTACACAGNGTGCATATTGGTGTCATCTAC	1544			
Db	323	ATGCAGCTACAAAGAAGGAAACACAGGATTGSCATCTCGCATCTTTGGCTGGGCAAGCAG	382			
Qy	1545	A5ACCTGCCGCCCTACTCCTTGAGCTATTGGGTGTGATCCTTAACATTATATCCCTTCAAGGGCT	1604			
Db	383	AGTGGTAAAGTCTTGTTGTACAAATGGAGGCCAATGCAATGCACAATCFCAGAAATGGTT	442			
Qy	1605	TTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCCTCCAAGAGGGTATCT---	1661			
Db	443	TCAGCGCNTGTATATGCGARCCCAGGAAATCACCTTGGAGTTGTGCAAGTTTCTTCTTGG	502			
Qy	1662	--CATTTAGSTTAATTCACAGGCCACAGACAATTTGCTGSAAGCTGCAAGGCTGGAGATGT	1719			
Db	503	ACAAATGGTGAAGCCASAGRCCTARCACAGARGATGGCTTCACACCAATTCGCAGTGGCITT	562			
Qy	1720	CGAAACTGTAAAAAAAACCTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCG	1779			

Db	563	TGCAACAAGGTCACGACCAAGTCGTTTTGCGCTCCTGCTAGAGAATGACACCAAGGAAAAG	622
Qy	1780	TCAGTCTCA-CACCACTTCATTTTGCAGCTGGGTATACAGAGTGTCCGTGGTGGAAATATC	1838
Db	623	TGCGTCTCCAGCTCTTCATATCGCGCCCGAAAAGACACGAAAGCGCGCGCCCTGC	682
Qy	1839	TGTCAGCATGGAGCTGATGTGCATGTAAGATAAAGGAGGCGTTGTACCTTTTGCACA	1898
Db	683	TGCTGCGAATGACACAAATGCAGATGCGAATCAAGAGTGGCTTCACTCCGCTCCACA	742
Qy	1899	ATGCATGTTCTTATGGACATTAATGAAGTTGCAGAACTCTTTGTTTAAACATGAGAGCAGTAG	1958
Db	743	TAGCTGCTCACTATGAAATATCAATGTGAGCCAGCTTGCTGTTTAAACCGARCGGCTGCTG	802
Qy	1959	TTAATGTAGCTGATTTATGGAAATTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAAAT	2018
Db	803	TGGAITTCACCGCAAGGAATGACATCTCTTTTACATGTTGCATCAAAAGAGGAAATG	862
Qy	2019	ATGAAATTTGCAAACTTCTGCTCCAGCATGTTGCAGACCCCTACCAAAAAAACAGGGATG	2078
Db	863	CAAAATATGGTAAAACTATTGCTCGATCGAGCAGCTAAAAATCGATGTCACCAAAACCGAGGATG	922
Qy	2079	GAAATACTCCTTTGATCTTG	2099
Db	923	GTCTGACACCACTGCACTGTG	943
RESULT 14			
US-09-947-199-3			
; Sequence 3, Application US/09947199			
; Patent No. US20020127684A1			
; GENERAL INFORMATION:			
; APPLICANT: Raju, Jeyaseelan			
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND US			
; TITLE OF INVENTION: THEREFOR			
; FILE REFERENCE: MNI-068CP2			
; CURRENT APPLICATION NUMBER: US/09/947,199			
; CURRENT FILING DATE: 2001-09-05			
; PRIOR APPLICATION NUMBER: 60/111,938			
; PRIOR FILING DATE: 1998-12-11			
; PRIOR APPLICATION NUMBER: 09/291,839			
; PRIOR FILING DATE: 1999-04-14			
; PRIOR APPLICATION NUMBER: 09/458,457			
; PRIOR FILING DATE: 1999-12-10			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 3			
; LENGTH: 2505			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)..(2505)			
US-09-947-199-3			
Query Match 2.0%; Score 74.2; DB 10; Length 2505;			
Best Local Similarity 55.1%; Pred. No. 2e-09;			
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps			
Qy	1787	ACACCACTTCATTTTGCAGCTGGGTATAACAGAGAGTGCCTGGTGGAAATATCTCTACAG	1846
Db	406	ACTGCCCTCCATATGCTACAATAGCTGCCACCTAGAGGCTGCTGATGCTGCTTGCAC	465
Qy	1847	CATGAGCTGATGTCATGCTAAAGATAAAGAGGCGCTTGTACCTTTCCACAATGCATGT	1906
Db	466	CATGAGCTAATGTCATATATTCAAGATGCAAGTTTTTTTCACTCCCATTCATATTGCAGCG	525
Qy	1907	TCTTATGGACATTATGAAGTTGTCAGAACTTCTTGTGTTAAACATGGAGCAGTAGTTAAATGTA	1966
Db	526	TACTATGGACATGNACAGGTAACTGCCCTTCTTTGGAATTTGGTGCTGATGTAATGTA	585
Qy	1967	GCTGATTTATGGAAATTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATT	2026

Db 586 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATT 645
Qy 2027 TGC AAACTTCTGCTCCAGCATGG 2049
Db 646 GC AAAACTCTTGATGGAAGAAGG 668

RESULT 15

US-09-947-199-1
; Sequence 1, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(2552)
US-09-947-199-1

Query Match 2.0%; Score 74.2; DB 10; Length 3025;
Best Local Similarity 55.1%; Pred. No. 2.3e-09;
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
Qy 1787 ACACCACATTCATTTTGCAGCTGGGTATACACAGAGTGCCGTTGGTGAATATCTGCTACAG 1846
Db 453 ACTGCCCTCCATATTGCTACATAGCTGGCCACCTAGAGGCTGCTGCTGTTCGAA 512
Qy 1847 CATGAGCTGATGCTGATGCTAAAGATAAAGAGGCCCTTGTAACCTTTGCACAAATGCATGT 1906
Db 513 CATGAGCTAATGCAATATCAAGATGCAGTTTTTTTCACTCCATTCATATTCGACGG 572
Qy 1907 TCTTATGGACATTTATGAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTAATGTA 1966
Db 573 TACTATGGACATGAACAGGTAACCTGCGCTCTTTTGAAATTTGGTCTGATGTAATGTA 632
Qy 1967 GCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAATTT 2026
Db 633 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATT 692
Qy 2027 TGC AAACTTCTGCTCCAGCATGG 2049
Db 693 GC AAAACTCTTGATGGAAGAAGG 715

Search completed: February 12, 2003, 03:31:24
Job time : 213.581 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 12:18:03 ; Search time 4936.65 seconds
(without alignments)
12456.682 Million cell updates/sec

Title: US-09-843-159B-1

Perfect score: 3797

Sequence: 1 ctttgagacactggatttc.....ccttgctgaaaaa 3797

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	784.2	20.7	838	9	AU140145	AU140145 AU140145
2	672.2	17.7	1128	13	BM457025	BM457025 AGENCOURT
3	649.2	17.1	659	14	R64714	R64714 EST54a09 WA
4	643.8	17.0	691	12	BG391376	BG391376 602417447
5	642.8	16.9	759	9	AL601027	AL601027 DKFZp313G
6	634.4	16.7	950	14	BQ885764	BQ885764 AGENCOURT

7	605	15.9	648	12	BG875327	RC1-BT062
8	593.2	15.6	622	10	BE220005	hvf65h03.x
9	584	15.4	616	10	BE467629	h267b11.x
10	570.4	15.0	956	12	BE910901	601661978
11	568	15.0	702	10	BB339554	BB339554
12	562.8	14.8	585	10	BE222311	hul10d01.x
13	557.4	14.7	595	10	BE504175	h586c11.x
14	549	14.5	553	14	BQ267594	ij90c10.y
15	544	14.3	1099	12	BG245969	602358786
16	538.6	14.2	562	10	BE504478	h257a02.x
17	534.4	14.1	996	12	BE585334	602102008
18	518.4	13.7	520	10	AW157349	au94h04.x
19	504	13.3	729	13	BG917245	602816521
20	503.4	13.3	553	12	BE721392	188553 MA
21	487.4	12.8	501	14	BM751146	K-EST0027
22	474.4	12.5	984	12	BF384484	602046448
23	462.2	12.2	600	12	BG807796	2072-17 M
24	459	12.1	460	9	AA307492	EST178376
25	456	12.0	748	14	BQ179839	UI-M-BW0-
26	450.8	11.9	645	10	BE373823	601227836
27	446.8	11.8	597	10	AW749617	RC3-BT050
28	437.4	11.5	1094	13	BM458871	AGENCOURT
29	431.8	11.4	441	10	AW208871	UI-H-B11-
30	430.6	11.3	580	13	BI064735	pgfln.pk0
31	420.4	11.1	709	13	BI152307	602917745
32	413.4	10.9	497	10	AW231078	uo70e07.y
33	405.8	10.7	469	10	BE503173	h284e01.x
34	401.8	10.6	598	9	AL680319	AL680319
35	399	10.5	411	9	AL711456	DKFZp6860
36	388	10.2	412	9	AL711468	DKFZp686P
37	380.4	10.0	384	9	AA244138	nc04g08.s
38	374.4	10.0	391	9	AA244137	nc04g08.r
39	377.6	9.9	458	12	BF021623	uy49g12.y
40	374.8	9.9	677	13	BJ038772	BJ038772
41	366.2	9.6	878	12	BF979977	602288547
42	362.8	9.6	819	13	BI558956	60324317
43	362	9.5	660	10	BB391900	BB391900
44	360	9.5	1114	13	BM464678	AGENCOURT
45	354.6	9.3	564	14	HI7748	ym40g08.r1

ALIGNMENTS

RESULT 1
AU140145
LOCUS AU140145 PLACE2 Homo sapiens cdna clone PLACE2000021 5', mRNA
DEFINITION AU140145 PLACE2 Homo sapiens cdna clone PLACE2000021 5', mRNA
ACCESSION AU140145
VERSION AU140145.1 GI:11001666
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 838)
AUTHORS Ota, T., Suzuki, Y., Saito, K., Ishii, S., Yamamoto, J., Sugiyama, T.,
Nishikawa, T., Nakamura, Y., Sugano, S., Masuho, Y., and Isogai, T.
TITLE HRI human cDNA project (Ota, T., Suzuki, Y., Saito, K., Ishii, S.,
Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S.,
Masuho, Y., Isogai, T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and

FEATURES	Helix Research Institute.	
	Location/Qualifiers	
SOURCE	1. 838	
	/organism="Homo sapiens"	
BASE COUNT	260 a 143 c 203 g 227 t	
	5 others	
ORIGIN	Query Match	
	Best Local Similarity 20.7%; Score 784.2; DB 9; Length 838;	
MATCHES	Matches 823; Conservative 0; Mismatches 13; Indels 4; Gaps 3;	
	QY 2715 CAGCCAGCAGTCTTGACACTTATCTGGAGAGTTTTCAGAACTGCTTCAGTAGTTAGTT 2774	
DB	1 CAGCCAGCAGTCTTGACACTTATCTGGAGAGTTTTCAGAACTGCTTCAGTAGTTAGTT 60	
	QY 2775 CAAGTGGAAACAGAGGCTGCTTCAGTTTGGAGAAAAGGAGGTTCCAGGAGTAGATTTTA 2834	
DB	61 CAAGTGGAAACAGAGGCTGCTTCAGTTTGGAGAAAAGGAGGTTCCAGGAGTAGATTTTA 120	
	QY 2835 GCATAACTCAATTCGPAACCAATCTTGGACTTCAGCACCTTAATGGATATATTTGAGAGAG 2894	
DB	121 GCATAACTCAATTCGPAACCAATCTTGGACTTCAGCACCTTAATGGATATATTTGAGAGAG 180	
	QY 2895 ACAGATCACTTTGGATGTATTAGTTGAGATGGGCACAGAGCTGAAGGAGATTGGAA 2954	
DB	181 ACAGATCACTTTGGATGTATTAGTTGAGATGGGCACAGAGCTGAAGGAGATTGGAA 240	
	QY 2955 TCAATGCTTATGGACATAGGCACAACTAATTAAGGAGCTCGAGAGACTTATCTCGGAC 3014	
DB	241 TCAATGCTTATGGACATAGGCACAACTAATTAAGGAGCTCGAGAGACTTATCTCGGAC 300	
	QY 3015 ACAAGGCTTTAACCCATATTTAACTTTGAACACCTCTCGTAGTGAACAAATCTTATAG 3074	
DB	301 ACAAGGCTTTAACCCATATTTAACTTTGAACACCTCTCGTAGTGAACAAATCTTATAG 360	
	QY 3075 ATCTGCTCTCATGATAGAGAGTTTCACTCTCTGGAGGAAGAGATGCAAGTACAGTTC 3134	
DB	361 ATCTGCTCTCATGATAGAGAGTTTCCGCTCTGGAGGAAGAGATGCAAGTACAGTTC 420	
	QY 3135 GAGGCACAGAGATGAGGTCATGCAAGTGGAACTCTTCAACAGATACAATATTTCAAGA 3194	
DB	421 GAGGCACAGAGATGAGGTCATGCAAGTGGAACTCTTCAACAGATACAATATTTCAAGA 480	
	QY 3195 TTCAGAGGTTTGTACAGAACTATGGAAAGATACACTCACCGGAGAAAGAGTTT 3254	
DB	481 TTCAGAGGTTTGTACAGAACTATGGAAAGATACACTCACCGGAGAAAGAGTTT 540	
	QY 3255 CTGAAGAAACACACCACTGCAATGAACGAATGCTATTTCATGGTCTCTCTTTGTGA 3314	
DB	541 CTGAAGAAACACACCACTGCAATGAACGAATGCTATTTCATGGTCTCTCTTTGTGA 600	
	QY 3315 ATGCAATATCCAAAGGCTTTGATGAAGGATGCGTACATAGTGATGTTGGAG 3374	
DB	601 ATGCAATATCCAAAGGCTTTGATGAAGGATGCGTACATAGTGATGTTGGAG 660	
	QY 3375 CTGGCAATTTATTTGCTGAAAACCTCTTCCAAAGCAATCAATATGATATGAAATGGAG 3434	
DB	661 CTGGCAATTTATTTGCTGAAAACCTCTTCCAAAGCAATCAATATGATATGAAATGGAN 719	
	QY 3435 GAGGTACTGGGTGCTCCAGTTCAAGACAGATCTTGTGTACATTTCCACAGCAGCTGC 3494	
DB	720 GAGGTACTGGGTGCTCCAGTTCAAGACAGATCTTGTGTACATTTCCACAGCAGCTGC 778	
	QY 3495 TCTTTTG--CCGGGTAAACCTTGGGAAAGTCTTCTCGAGTTTCAGTGCAGTGAATGAAATGCG 3552	
DB	779 TCTTTTGCCGGGTAAACCTTGGGAAAGTCTTCTCGAGTTTCAGTGCAGTGAATGAAATGCG 838	
RESULT 2	BM457025	
	AGENCOURT_6406851 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583257	
LOCUS	5', mRNA sequence.	
	BM457025	
ACCESSION	BM457025	
	BM457025.1 GI:18506065	
VERSION	EST.	
	human.	
KEYWORDS	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
SOURCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	1 (bases 1 to 1128)	
ORGANISM	NIH-MGC http://mgc.nci.nih.gov/.	
	National Institutes of Health, Mammalian Gene Collection (MGC)	
REFERENCE	Unpublished (1999)	
	Contact: Robert Strausberg, Ph.D.	
AUTHORS	Email: cgapbs-r@mail.nih.gov	
	Tissue procurement: ATCC	
TITLE	cDNA Library Preparation: Life Technologies, Inc.	
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
JOURNAL	DNA Sequencing by: Agencourt Bioscience Corporation	
	Clone distribution: MGC clone distribution information can be	
COMMENT	found through the I.M.A.G.E. Consortium/LLNL at:	
	http://image.llnl.gov	
FEATURES	Plate: LLAM12346 row: a column: 18	
	High quality sequence stop: 507.	
source	Location/Qualifiers	
	1. 1128	
BASE COUNT	328 a 297 c 232 g 271 t	
	Note: this is a NIH_MGC Library.	
ORIGIN	Query Match	
	Best Local Similarity 17.7%; Score 672.2; DB 13; Length 1128;	
MATCHES	Matches 724; Conservative 0; Mismatches 13; Indels 4; Gaps 4;	
	QY 1957 AGTTAATGTAGCTGATTTATGGAATTTACACCTTTACATGAGCAGCAGCAAGGAA 2016	
DB	1 AGTTAATGTAGCTGATTTATGGAATTTACACCTTTACATGAGCAGCAGCAAGGAA 60	
	QY 2017 ATATGAAATTTGCAAACTTCTGCTCCAGCATGTCAGACCTTACCACCAAAAAACAGGGA 2076	
DB	61 ATATGAAATTTGCAAACTTCTGCTCCAGCATGTCAGACCTTACCACCAAAAAACAGGGA 120	
	QY 2077 TGAATATCTCTTGGATCTTGTAAAGATGGAGATACAGATATTCAAGATCTGCTTAG 2136	
DB	121 TGAATATCTCTTGGATCTTGTAAAGATGGAGATACAGATATTCAAGATCTGCTTAG 180	
	QY 2137 GGGAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTAGCCAGAGTGAAGAAGTT 2196	
DB	181 GGGAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTAGCCAGAGTGAAGAAGTT 240	
	QY 2197 GTCCTCTCTGATATAATGTAATGTCGCCGATACCCAGGCAGACATTCACACCTTTTACA 2256	
DB	241 GTCCTCTCTGATATAATGTAATGTCGCCGATACCCAGGCAGACATTCACACCTTTTACA 300	
	QY 2257 TTATGAGCTGTTTATTAATTAATTTAGAAAGTTCAGAGTATTTGTTTACAACAGGAGCTGA 2316	
DB	301 TTATGAGCTGTTTATTAATTAATTTAGAAAGTTCAGAGTATTTGTTTACAACAGGAGCTGA 360	
	QY 2317 TGTGAATGCCCAGACAAAGGAGGACTTATCTTTACATATATGACAGATCTTACGGGCA 2376	


```

Db 361 TGTGAATGCCCAAGACAAAGAGGAGCTATTTCCTTTACATATATGCGAGCATCTTACGGGCA 420
QY 2377 TGTAGATGTAGCAGCTCTACTATAAAGCTATATGATGTGTCAATGCCCGGACAAATG 2436
Db 421 TGTAGATGTAGCAGCTCTACTATAAAGCTATATGATGTGTCAATGCCCGGACAAATG 480
QY 2437 GGCCTTTCACACCTTTGCGAGGAGCCGCAAAAGGAGCAGACACAGCTTTGTGCTTTGTT 2496
Db 481 GGCCTTTCACACCTTTGCGAGGAGCCGCAAAAGGAGCAGACACAGCTTTGTGCTTTGTT 540
QY 2497 GCTAGCCATGAGCTGACCCGACTTTAAATATCAGGAAGGACAAACA-CCTTTAGATT 2555
Db 541 GCTAGCCATGAGCTGACCCGACTTTAAATATCAGGAAGGACAAACACCTTTTACATT 600
QY 2556 TAGTTTTCAGCGGATGATGTCAGCG-CTCTTCTGACAGCAGCATGCCCCCATCTGCTCTG 2614
Db 601 TAGTTTTCAGCGGATGATGTCAGCGCTCTCTGACACCGCATGGCCCCATCTGCTCTG 660
QY 2615 CCCTCTTTTACAAGCCTCAAGTGTCTCAATGGTG-TGAGAAAGCCAGAGCCACTGCGAGA 2673
Db 661 CCCTCTTTGTCAGCGCTCAAGCGCTCCATGGTGTGAGAAAGCCAGCGCCACTGCGAGA 720
QY 2674 TG-CTCTCTCTTCAGTCCAT 2693
Db 721 TGCTTTCTCTTCAGGTTCTCT 741

RESULT 3
R64714
LOCUS
DEFINITION
EST54409 WATM1 Homo sapiens cDNA clone 54a09 similar to BLASTX
homology with sp|Q01485|ANKC_HUMAN ANKYRIN, BRAIN VARIANT 2, mRNA
sequence.
R64714
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Bouillaud,F.
1 (bases 1 to 659)
Study of expressed sequences tags in adipose tissue 1995
Unpublished (1995)
Contact: Frederic Bouillaud
Centre de Recherche sur l'Endocrinologie moleculaire et le
Developpement
CNRS
9, Rue Jules Hertz, Meudon Bellevue, 92190 France
Tel: 33 1 45 07 52 87
Fax: 33 1 45 07 58 90
Email: bouillaud@infobiogen.fr
automatic cycle sequencing of PCR amplified insert, open reading
frame 3. 659.
FEATURES
source
location/Qualifiers
1..659
/organism="Homo sapiens"
/strain="Caucasian"
/db_xref="taxon:9606"
/clone="54a09"
/clone_lib="WATM1"
/note="Vector: lambda gtl; Site_1: EcoRI; White adipose
tissue, sub cutaneous, adult, female. Purification of
polya mRNA, first strand priming with random
oligonucleotides. Ligation of adaptors EcoRI/NotI, cloning
in EcoRI site of lambda gtl."
200 a 128 c 154 g 172 t 5 others
BASE COUNT
ORIGIN
Query Match 17.1%; Score 649.2; DB 14; Length 659;
Best Local Similarity 98.6%; Pred. No. 1.1e-147;
Matches 649; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1416 TGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAACATGAAG 1475
Db 1 TGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAACATGAAG 60
QY 1476 CAAAGGTTAATGCTCTCGATAAATCTTGGTCAGACTTCTTACACAGAGCTGCATATTGTG 1535
Db 61 CAAAGGTTAATGCTCTCGATAAATCTTGGTCAGACTTCTTACACAGAGCTGCATATTGTG 120
QY 1536 GTCATCTACAAACCTGCGCCTACTCCTGAGCTATGGTGTGATCCCTAACATATATATCCC 1595
Db 121 GTCATCTACAAACCTGCGCCTACTCCTGAGCTATGGTGTGATCCCTAACATATATATCCC 180
QY 1596 TTCAGGCTTTTACTGCTTTACAGATGGAAATGAAAATGTACAGCAACTCTCTCCAAGAGG 1655
Db 181 TTCAGGCTTTTACTGCTTTACAGATGGAAATGAAAATGTACAGCAACTCTCTCCAAGAGG 240
QY 1656 GTATCTCATTTAGTAAATTCAGAGGCAGACAGACAATTTGCTGGAGCTGCAAGGCTGGAG 1715
Db 241 GTATCTCATTTAGTAAATTCAGAGGCAGACAGACAATTTGCTGGAGCTSCAAGGCTGGAG 300
QY 1716 ATGTCGAAACTGTAAAAAACTGTGTACTGTTTCAGACTGTCAACTGCAGAGACATTGAAG 1775
Db 301 ATGTCGAAACTGTAAAAAACTGTGTACTGTTTCAGACTGTCAACTGCAGAGACATTGAAG 360
QY 1776 GGCCTCAGTCTACACCACTTTCATTTTGCAGCTGGGTATACAGAGTGTCCGTGTGGAAT 1835
Db 361 GGCCTCAGTCTACACCACTTTCATTTTGCAGCTGGGTATACAGAGTGTCCGTGTGGAAT 420
QY 1836 ATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGCTTGTACCTTTGC 1895
Db 421 ATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGCTTGTACCTTTGC 480
QY 1896 ACAATGATGTTCTTATGAGCATATATCAAGTGCAGAACTTCTTGTAAACATGGAGCAG 1955
Db 481 ACAATGATGTTCTTATGAGCATATATCAAGTGCAGAACTTCTTGTAAACATGGAGCAG 540
QY 1956 TAGTTAATGTAGTGTATTTATGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAA 2015
Db 541 TAGTTAATGTAGTGTATTTATGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAA 600
QY 2016 AATATGAAATTTGCAAACTTCTGCTCCAGCATGTCGAGACCCCTACCAAAAAAACAG 2073
Db 601 AATATGAAATTTGCAAACTTCTGCTCCAGCATGTCGAGACCCCTACCAAAAAAACAG 658

RESULT 4
BG391376
LOCUS
DEFINITION
602417447F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4537138 5',
mRNA sequence.
BG391376
ACCESSION
BG391376.1 GI:13284824
VERSION
BG391376.1
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 691)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMI0461 row: m column: 11
High quality sequence stop: 684.
Location/Qualifiers

```

FEATURES

source	1. .691 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4537138" /clone_lib="NIH_MGC_92" /tissue_type="embryonal carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."	218 a 122 c 165 g 186 t	
BASE COUNT			
ORIGIN			
Query Match	17.0%;	Score 643.8;	DB 12; Length 691;
Best Local Similarity	99.3%;	Pred. No. 2.3e-146;	
Matches 678;	Conservative 0;	Mismatches 2;	Indels 3; Gaps 3;
Qy 1706	AAGCGTGGAGATGTCGAAACTGTAAAAAAA	CTGTACTGTTCCAGAGTGTCAAACTGCAGA	1765
Db 10	AAGCGTGGAGATGTCGAAACTGTAAAAAAA	CTGTACTGTTCCAGAGTGTCAAACTGCAGA	69
Qy 1766	GACATTGAAGGGCTCAGTCTACACCACTTCA	TTTTCGACGCTGGGTATACAGAGTGTCC	1825
Db 70	GACATTGAAGGGCTCAGTCTACACCACTTCA	TTTTCGACGCTGGGTATACAGAGTGTCC	129
Qy 1826	GTGCTGGAATATCGCTACAGCATGGAGCTG	ATGCTCATCTAAAGATAAAGGAGGCTT	1885
Db 130	GTGCTGGAATATCGCTACAGCATGGAGCTG	ATGCTCATCTAAAGATAAAGGAGGCTT	189
Qy 1886	GTACCTTTGGCACAATGCATGTTCTTATG	GCATTATGAAGTTCGACAACTTCTGTTAAA	1945
Db 190	GTACCTTTGGCACAATGCATGTTCTTATG	GCATTATGAAGTTCGACAACTTCTGTTAAA	249
Qy 1946	CATGGACAGTAGTTAATGTAGCTGATTATG	GAATTTACACCTTTACATGAAGCAGCA	2005
Db 250	CATGGACAGTAGTTAATGTAGCTGATTATG	GAATTTACACCTTTACATGAAGCAGCA	309
Qy 2006	GCAAAAGGAAATATGAAATTTGCAAACTTC	TGCTCCAGCA-TGGTGCAGACCCCTACAA	2064
Db 310	GCAAAAGGAAATATGAAATTTGCAAACTTC	TGCTCCAGCACTGGTGCAGACCCCTACAA	369
Qy 2065	AAAAACAGGATGGAATACTCCTTTGGATCT	TGTTTAAAGATGGAGATACAGATATTC	2124
Db 370	AAAAACAGGATGGAATACTCCTTTGGATCT	TGTTTAAAGATGGAGATACAGATATTC	429
Qy 2125	AGATCTGCTTAGGGAGATGCAGCTTTGCT	AGATGCTGCAAGAGGGTGTGTAGCCAG	2184
Db 430	AGATCTGCTTAGGGAGATGCAGCTTTGCT	AGATGCTGCAAGAGGGTGTGTAGCCAG	489
Qy 2185	AGTCAAGAGTGTGCTCTCCTGATA-ATG	TAAATGCCCGCATACCCAGGACAGAT	2243
Db 490	AGTCAAGAGTGTGCTCTCCTGATAATG	TAAATGCCCGCATACCCAGGACAGAT	549
Qy 2244	CAACACCTTTACATTTAGCAGCTGGTTAT	ATATTAATTTAGAAATTCAGAGTATTTG	2303
Db 550	CAACACCTTTACATTTAGCAGCTGGTTAT	ATATTAATTTAGAAATTCAGAGTATTTG	608
Qy 2304	AACACGAGCTGATGTGAATGCCCAAGACA	AGAGGAGCTTATTCCCTTTACATAATGC	2363
Db 609	AACACGAGCTGATGTGAATGCCCAAGACA	AGAGGAGCTTATTCCCTTTACATAATGC	668
Qy 2364	CATCTTACGGGCATGTAGATGTA	2386	
Db 669	CATCTTACGGGCATGTAGATGTA	691	
RESULT 5			
AL601027	AL601027	759 bp	linear EST 14-AUG-2001
LOCUS	DRFZp313G2239_r1	313 (synonym: hlcc2)	Homo sapiens cdna clone
DEFINITION	DRFZp313G2239 5', mRNA sequence.		

```
QY 2854 GAATCTTGAGCTGAGCACCTAAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGT 2913
Db 552 GAACTCTTGAGCTGAGCACCTAAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGT 611
QY 2914 ATTACTTCAGATGGGCGACAGGAGCTGAAGGAGATTGGAATCAATCCTTATGACATAG 2973
Db 612 ANAGTTGAGATGGGCGACAGGAGCTGAAGGAGATTGGAATCAATCCTTATGACATAG 671
QY 2974 GCACAACTAATTTAA-GGAGTCGAGAGACTTATCTCCGGCAACAA-GGTCTTTAAACCCA 3031
Db 672 GCACAACTAATTTAAAGGAGTCGAGAGACTTATCTCCGGCAACAAAGGGCTTTAAACCCA 731
QY 3032 TATTTAACTTTGAACACCTCTGGTAGTG 3059
Db 732 TATTTAACTTTGAACACCTCTGGNAGG 759

RESULT 6
BQ885764
LOCUS
DEFINITION
AGENCOURT_8746924 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6392888
5', mRNA sequence.
ACCESSION
BQ885764
VERSION
BQ885764.1 GI:22277782
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 950)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAM13883 row: h column: 09
High quality sequence stop: 641.
FEATURES
Location/Qualifiers
1..950
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6392888"
/clone_lib="NIH_MGC_130"
/lab_host="DH10B (phage-resistant)"
/note="Organ: otcysts; Vector: pCMV-SPORT6.1.cdbb;
Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally.
Primer: Oligo dt. Average insert size 1.95 kb.
Constructed by ResGen, Invitrogen Corp. Note: this is a
NIH_MGC Library."
BASE COUNT 280 a 195 c 226 g 245 t 4 others
ORIGIN

Query Match 16.7%; Score 634.4; DB 14; Length 950;
Best Local Similarity 86.4%; Pred. No. 5,2e-144;
Matches 759; Conservative 0; Mismatches 111; Indels 8; Gaps 5;

QY 2481 AGCTTTGCTTTGTTGCTAGCCCATGAGCTGACCCGACTTTAAAAATCAGGAAGGAC 2540
Db 1 AGCTTTGCTTTGTTGTTGCTAGCCCATGAGCTGACCCGACTTTAAAAATCAGGAAGGAC 59
QY 2541 AAACACCTTTAGATTTTACGGGGATGATGTCAGGGCTCTTCTGACAGCAGCCATGC 2600
Db 60 AAACACCTTTAGATTTTACGGGGATGATGTCAGGGCTCTTCTGACAGCAGCCATGC 119
QY 2601 CCCCATCTGCTGCTGCTCTTTTACAAGCCTCAAGTCTCAATGGTGTGAGAAGCCAG 2660
```

```
Db 120 CCCCCCTCTGCTCGCCTACGTGCTACAAAACCTCAAGTGCCTCAGTGGCGTCAAGGGCCCCG 179
QY 2661 GAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTACGCCCATCAAGCCCTTCTTCGACCCA 2720
Db 180 GAGCCACTGCAGATGCTCTCTCTTCAGGTCCGCTCCAGGCCCATCCAGCCTCTCTCGACCCA 239
QY 2721 GCAGTCTTCACAACCTTATCTCGGAGTTCAGAACTGCTTTCAGTAGTTCAGTTCAGTGC 2780
Db 240 GCAGCCTCGACAACCTTATCTCGGAGCTTCTCGCAACTGTCGCGAGTGGTTCAGTTCAGTGC 299
QY 2781 GAACAGAGGTGCTTCCAGTTTCGAGAAAGAGGTTCCAGGAGTAGATTTTATGACATAA 2840
Db 300 CAGCAGAAGGTGCTACTGGTTTGCAAAGAAAGAGGATTTCAGGAATCGATTTTATGATATA 359
QY 2841 CTCAATTTCGTAAGGAACTCTTTGGACTTCAGCACCTTAATGGATATATTTGAGAGAGACAGA 2900
Db 360 CTGAGTTTCATAAGGAACCTTTGGACTTCAGCACCTTGATGATATATTTGAGAGAGACAGA 419
QY 2901 TCACCTTTGGATGTTAGTTGAGATGGGCGACCAAGGAGCTGAAGGAGATTGGAATCAATGC 2960
Db 420 TTACCTTAGATGCTCTCTCTTAACTCTAAACAACTCTGTTAGTGGAACTTCTCATATGATCTGT 599
QY 2961 CTTATGAGATAGCGACAAACTAATTTAAAGAGTTCGAGAGACTTATCTCCGGCAACAAAG 3020
Db 480 CCTATGAGATAGCGACAACTAATTTAAAGAGTTCGAGAGACTTATCTCCGGCAACAAAG 539
QY 3021 GTCTTAAACCATATTTAACTTTGAACACCTCTGTTAGTGGGCGACCAAGGAGATTGGAATCAATGC 3080
Db 540 GTCTTAAACCATATTTAACTTTGAACACCTCTGTTAGTGGGCGACCAAGGAGATTGGAATCAATGC 599
QY 3081 CTCTCTGATGATAAGAGTTCACGTCTGTGGAGGAAGAGATGCAAACTACAGTTCGAGAGC 3140
Db 600 CTCTCTGATGATAAGAGTTCACGTCTGTGGAGGAAGAGATGCAAACTACAGTTCGAGAGC 659
QY 3141 ACAGAGATGAGGTGATGTCAGGTGGAATCTTCAACAGATACATAATTTCAAGATTTCAGA 3200
Db 660 ACAGAGATGAGGTGATGTCAGGTGGAATCTTCAACAGATACATAATTTCAAGATTTCAGA 719
QY 3201 AGGTTTGTACAGAACTATGGAAGATACACTACCCGG-AGAAAGAGTTCCTCGAA 3259
Db 720 AGGTTTGTACAGAACTATGGAAGATACACTACCCGGAGATATACACCCGGNAGAAAGAGTTCCTCGAA 779
QY 3260 GAAACACCAACCATGC-CAATGAACGAATGCTATTTCATGGGTCT-CTTTCATGGTGA 3315
Db 780 GAAACACCAACCATGC-CAATGAACGAATGCTATTTCATGGGTCT-CTTTCATGGTGA 839
QY 3316 TGCAATATCCACAAGG--CTTTGATGAAGGCATGC 3351
Db 840 TGGGATTATCCATAAAGGGCTTTTGTGATGAAGGCATGC 877

RESULT 7
BQ875327
LOCUS
DEFINITION
RC1-BT0623-120200-011-d04 BT0623 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BQ875327
VERSION
BQ875327.1 GI:14252306
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 648)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL
```

MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=st2-RC1-BT0623-120
200-011-d04&t3=2000-02-12&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 4
High quality sequence stop: 648.
Location/Qualifiers
1..648
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT0623"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
200 a 118 c 157 g 172 t 1 others
BASE COUNT 200 a 118 c 157 g 172 t 1 others
ORIGIN
Query Match 15.9%; Score 605; DB 12; Length 648;
Best Local Similarity 98.7%; Pred. No. 7.1e-137;
Matches 620; Conservative 0; Mismatches 6; Indels 2; Gaps 1;
QY 1644 TCCTCCAGAGGGTATCTCATTTAGTAAATTCAGAGGCAGACAGACAATTCGTCGAGAGCTG 1703
DB 23 TCCTCAAGAGGGTATCTCATTTAGG--ATTCCAGAGGCAGACAGACAATTCGTCGAGAGCTG 80
QY 1704 CAAAGCTCGAGATGTCGAAACTGTAAAAAACTGTGTACTGTCAGAGTGTCAACTGCA 1763
DB 81 CAAAGCTCGAGATGTCGAACTGTAAAAAACTGTGTACTGTCAGAGTGTCAACTGCA 140
QY 1764 GAGACATTGAAGGGCGTCACTACACCACTTCATTTTCGACGTGGGTATACAGAGTGT 1823
DB 141 GAGACATTGAAGGGCGTCACTACACCCNTTCATTTTCAGCTGGGTATACAGAGTGT 200
QY 1824 CCGTGGTGAATATCTGTACAGCATGGAGCTGATGCTGATCTAAAGATAAAGAGGCC 1883
DB 201 CCGTGGTGAATATCTGTACAGCATGGAGCTGATGCTGATCTAAAGATAAAGAGGCC 260
QY 1884 TTGTACCTTTGCACATGTCATGTTCTTATGACATTATGAAGTTGACAGAACTTCCTTTGTTA 1943
DB 261 TTGTACCTTTGCACATGTCATGTTCTTATGACATTATGAAGTTGACAGAACTTCCTTTGTTA 320
QY 1944 ACATGGAGCAGTAGTAAATGATGATGATTTATGAAATTTACACTTTACATGAAGCAG 2003
DB 321 AACATGGAGCAGTAGTAAATGATGATGATTTATGAAATTTACACTTTACATGAAGCAG 380
QY 2004 CAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGTCGAGACCCTACCA 2063
DB 381 CAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGTCGAGACCCTACCA 440
QY 2064 AAAAAACAGGGATGGAATACTCCCTTTTGGATCTTTGTTAAAGATGGAGATACAGATATTC 2123
DB 441 AAAAAACAGGGATGGAATACTCCCTTTTGGATCTTTGTTAAAGATGGAGATACAGATATTC 500
QY 2124 AGATCTGCTTAGGGAGATGACAGCTTCCTAGATGCTGCCAAGAGGGTTGTTAGCCA 2183
DB 501 AGATCTGCTTAGGGAGATGACAGCTTCCTAGATGCTGCCAAGAGGGTTGTTAGCCA 560
QY 2184 GAGTGAAGAAGTGTCTTCTCTCTGATATGATGTAATTCGCCGCATACCCCAAGGCAGACATT 2243

DB 561 GAGTGAAGAAGTGTCTTCTCTGATATGTAATTCGCGCATATCCCAAGGCAGACATT 620
QY 2244 CAACACCTTTACATTTAGCAGCTGGTTA 2271
DB 621 CAACACCTTTACATATAGCAGCTGGTTA 648
RESULT 8
BE220005/c
LOCUS BE220005 622 bp mRNA linear EST 03-JUL-2000
DEFINITION hv5h03.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3178325 3'
similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED
ADP-RIBOSE POLYMERASE. ; mRNA sequence.
ACCESSION BE220005
VERSION BE220005.1 GI:8907323
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 622)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: "40UP from Gibco
High quality sequence stop: 474.
Location/Qualifiers
1..622
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3178325"
/clone_lib="NCI_CGAP_Lu24"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pT7T3D-pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI_CGAP_Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneIDs
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 179 a 130 c 128 g 185 t
ORIGIN
Query Match 15.6%; Score 593.2; DB 10; Length 622;
Best Local Similarity 97.1%; Pred. No. 5.4e-134;
Matches 604; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 3164 GGAATCTTCAACAGATACAAATATTCACAGATTCAGAAGGTTTGTAAACAGAACTATGG 3223
DB 622 GCAATCTTCCACAGACTACTATATTCACAGATTCACAGAGTCTGTACACATCTACTATCG 563
QY 3224 GAAAGATACATCCCGGAGAAAAGAGTTCTCTGAAGAAAACACCAACCATGCAATGAA 3283
DB 562 CAAAGATACATCCACGGCAGAAAACAGAGTTCTTGAAGAAAACCCCAACCATGCAATGAA 503
QY 3284 CGAATGCTATTTTCATGGGTCTCCTTTTGTGAATGCAATTTATCCAAAGAGCTTTGTATGA 3343
DB 502 CGAATGCTATTTTCATGGGTCTCCTTTTGTGAATGCAATTTATCCAAAGAGCTTTGTATGA 443

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue procurement: Gilbert Smith, Ph.D.
 cDNA library Preparation: Life Technologies, Inc.
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Place: LLAM9128 row: k column: 20
 High quality sequence stop: 672.
FEATURES Location/Qualifiers
 1..956
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:3962251"
 /clone_lib="NCI_CGAP_Maml"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="pDH108"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Sall; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 292 a 183 c 238 g 242 t
ORIGIN
 Query Match 15.0%; Score 570.4; DB 12; Length 956;
 Best Local Similarity 85.2%; Pred. No. 2.4e-128;
 Matches 673; Conservative 0; Mismatches 111; Indels 6; Gaps 3;
 QY 2828 GATTTTAGCACTCAATTCGTAAGGAATCTTGGACTTGGACACCTAATGGATATATTT 2887
 DB 1 GATTTTAGTATCACTCAGTTCATGAAGAACCTTGGACTTGGACACTTGAATGATATATTT 60
 QY 2888 GAGAGAGAACAGCATCTTTGGATGATTAGTTAGATGGGGCACAAGGAGCTGAAGAG 2947
 DB 61 GAGAGAGAACAGATTACCTTAGATGCTCTAGTTGAATGGCCACAAGCACTGAAAGAG 120
 QY 2948 ATTGGGAATCAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCAGACACTTATC 3007
 DB 121 ATTGGGAATCAATGCCCTATGGACATAGGCACAACTGATTAAGGAGCTTGAAGGCTGATC 180
 QY 3008 TCCGGACAACAAGGCTTTAACCCATATTTAACTTTGAACACCTCTGCTAGTGGAACAATT 3067
 DB 181 TCTGGACACACAGGCTTTAATCCATATTTAACTCTAAACACTCTGCTAGTGGAACAATT 240
 QY 3068 CTTATAGATCTCTCTCATGATAAAGAGTTTACGTTCTGTGGAGGAAGAGATGCAAGT 3127
 DB 241 CTCATAGATCTGCTCTCATGATAAAGAAATTTACGTTCTGTGGAGGAAGAGATGCAAGT 300
 QY 3128 ACAGTTCGAGACACAGATGAGGTCATGAGGTGGAATCTTCAACAGATACATATTT 3187
 DB 301 ACTGTGCGGGGACAGAGATGGTGTACGACGGCGGCTTTCAACAGATACACATTT 360
 QY 3188 CTCAGATTTCAGAAGGTTTGTAAACAAGAACTATGGAAGAATACACTCACCGGAGAAAA 3247
 DB 361 CTCAGATTTCAGAAGGTTTGTAAACAAGAAATTTGTGGAAGAATATACACACCGGAGAAAA 420
 QY 3248 GAAGTTTCGAGAAACACACACCAATGCCATGAACGAATGCTATTTTCATGGGTCTCT 3307
 DB 421 GAAGTTTCGAGAAACACACACCAATGCCATGAACGAATGCTATTTTCATGGGTCTCT 480
 QY 3308 TTTGTGAATGCAATATCCACAAGGCTTTGATGAAGGCAATCGCTACATAGGTGATG 3367
 DB 481 TTTGTGAATGCAATATCCATAGGCTTTGATGAAGGCAATCGCTACATAGGTGATG 540
 QY 3368 TTTGGAGCTGGCATTTATTTTCTGCTGAAAACTCTTCCAAAAGCAATCAATATGATGGA 3427

DB 541 TTTGGAGCTGGAAATTTATTTTCTGCTGAGAACTCTTCCAAAAGCAATCAATATGTGTATGGA 600
 QY 3428 ATTGGAGGAGG---TACTGGGTCTCCAGTTCACAAAAGACAGATCTTGTACATTTGGCCAC 3484
 DB 601 ACTTGGAGGTGGCACCCGGGATGTCCAATTCACAAAGACAGATCGTGTACCTTTGTCCAC 660
 QY 3485 AGCAGCTCTCTTTTGGCGGGTAACTTTGGGAAAGTCTTTCTCTGAGTTCAGTGAATG 3544
 DB 661 AGCAGCTCTGTTTCTGTCGAGT-ACCTTGGGCAAGTCTTTCTTCTGCGTTCACAGCATTTG 719
 QY 3545 ARAATGGCACAATCTCTCTCCAGTTCATCACTCACTCACTGG--TAGCCCAAGTGTAAATG 3602
 DB 720 ACAATGGACATTTCTCTCCGCGCCATCAATTCGGGCCACTGGCCCGGCCAGTGTGAATG 779
 QY 3603 GCCTAGCATT 3612
 DB 780 GCCTAGCATT 789
 RESULT 11
 BB339554
 LOCUS
 DEFINITION BB339554 702 bp mRNA linear EST 24-OCT-2001
 musculus cDNA clone B930018021 3', mRNA sequence.
 ACCESSION BB339554
 VERSION BB339554.2 GI:16404413
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 702)
 Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
 Hiramoto,K., Hori,P., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okada,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,F., Suzuki,H.,
 Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
 Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 Unpublished (2001)
 On Jul 11, 2000 this sequence version replaced gi:9048317.
 Contact: Yoshhide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL:<http://genome.gsc.riken.go.jp/>
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,
 M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,
 S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
 Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,
 Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Konno,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,
 K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
 Hayashizaki,Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES

source
1..702
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="B930018021"
/clone_lib="RIKEN full-length enriched, 10 days neonate cerebellum"
/sex="mixed"
/tissue_type="cerebellum"
/dev_stage="10 days neonate"
/lab_host="DH10B"
/note="Site:1: Sali; Site:2: BamHI: cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTCCAGTTAAATAATTAATCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"

BASE COUNT 202 a 160 c 173 g 167 t
ORIGIN

Query Match 15.0%; Score 568; DB 10; Length 702;

Best Local Similarity 89.7%; Pred. No. 8.2e-128;

Matches 621; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

Qy 1313 TGTGCTGCTGATCCATATCCCAAGAAGCAATATGTGACCTGCTGAAGAAA 1372

Db 12 TGTGCTGCTGATCT-CATATCCCAAGAAGCAGATATGTGAACGTGCTCAGAAAG 70

Qy 1373 GGAGCAACATCAATGAAAGACTAAAGAAATCTTGACTCTCTGCAGTGGCATCTGAG 1432

Db 71 GGAGCAACACCAAGAAAGACTAAAGAAATCTTGACTCTCTGCAGTGGCATCTGAG 130

Qy 1433 AAGCTCATATGATGTTGTTGAAGTAGTGTGAAACATGAAGCAAGGTTAATGCTCTG 1492

Db 131 AACGCTCAATGATGTTGTTGAAGTAGTGTGAAACATGAAGCAAGGTTAATGCTCTG 190

Qy 1493 GATATCTTGTGACACTCTCTACAGAGCTGCATATTTGCTCATCTCAACACCTGC 1552

Db 191 GATAGCTTGGACAGAGCTCATACAGAGCTGCACACTGTGTCACCTGCAGACCTGC 250

Qy 1553 CGCCTACTCTGAGCTATGGGTGTGATCCTTAACATTTATATCCCTTCAGGGGCTTACTGCT 1612

Db 251 CGCCTGCTCTGAGCTATGGGTGTGATCCTTAACATTTATATCCCTTCAGGGGCTTACTGCT 310

Qy 1613 TTACAGATGGGAATGAAATGTACAGCAACTCTCCCAAGAGGGTATCTCATTTAGGTAAT 1672

Db 311 TTGCAATGGGAATGAAATGTGAGCAGCTGCTCCCAAGAGGGGCGCTCACTAGGTCAC 370

Qy 1673 TCAGAGGCAGACAGCAATTCCTGAAGCTGCAAGGGCTGGAGATGTCGAACCTGTAAA 1732

Db 371 TCAGAGGGGAGACAGACAGCTGTGAAGCTGCAAGGGCTGGAGATGTCGAACCTGTAAA 430

Qy 1733 AAATGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGCTCACTACACCA 1792

Db 431 AAACCTGCTGCTTTCAGAGTGTCAACATGCAGAGACATTGAAGGACGCCAGTCAACCCCA 490
Qy 1793 CTTTCATTTTTCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGA 1852
Db 491 CTCCTACTTTCAGCTGGGTATACAGAGTGTCTGTGGTGAATATCTGCTGCAACATGA 550
Qy 1853 GCTGATGTCATGCTAAAGATAAAGGAGCGCTTGTACCTTTTCACAAATGCATGTTCTTAT 1912
Db 551 GCTGATGTCATGCTTAAACACAAAGGAGGTCTTGTACCTTTACACAAATGCATGTTCTTAT 610
Qy 1913 GGACATTTATGAAGTTGTCAGAACTCTTGTAAACATGGAGCAGTAGTTAATGTAAGTGTAT 1972
Db 611 GGACACTATGAAGTTGTCAGAACTCTTGTGCAAGCATGGAGCAGTAGTCAATGTCGCTGAC 670
Qy 1973 TTATGGAATTTACACCTTTTACATGAAGCAGC 2004
Db 671 TTGTGGAAGTTTACACCTTTTACATGAAGCTGC 702

RESULT 12

BE222311/c

LOCUS

DEFINITION

ADP-RIBOSE POLYMERASE. ;, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE222311 585 bp mRNA linear EST 03-JUL-2000
hu10d01.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3166177 3'
similar to TR:095271 095271 TRF1-INTERACTING ANKYRIN-RELATED
ADP-RIBOSE POLYMERASE. ;, mRNA sequence.

BE222311 1 GI:8909629

EST.

human.

Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 585)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cqabps-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov

Seq primer: -400P from Gibco

High quality sequence stop: 476.

Location/Qualifiers

source

1..585

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3166177"

/clone_lib="NCI_CGAP_Lu24"

/tissue_type="carcinoid"

/lab_host="DH10B"

/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; plasmid DNA from the normalized library NCI_CGAP_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction.

The driver was PCR-amplified cDNAs from a pool of 5,000

clones made from the same library (cloneIDs

1414920-1417991 and 1520904-1522439). Subtraction by Bento

Soares and M. Fatima Bonaldo.

BASE COUNT 166 a 126 c 112 g 176 t

ORIGIN

Query Match 14.8%; Score 562.8; DB 10; Length 585;

Best Local Similarity 98.5%; Pred. No. 1.4e-126;

Matches 576; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

Query Match	14.5%;	Score 549;	DB 14;	Length 553;
Best Local Similarity	100.0%;	Pred. No. 3.3e-123;		
Matches 549;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 3241	GAGAAAGAAGTTTCTGAAGAAAACCCACACCATGCCAATGAACGAATGCTATTTCATGG	3300		
QY 3301	GTCTCCTTTTGTGAATGCAATTATCCACAAGGCTTTTGATGAAGGCATGCGTACATAGG	3360		
QY 3361	TGSTATGTTTGAGCTGGCATTATTTTGCTGAAAACCTTCCAAAAGCAATCAATATGT	3420		
QY 3421	ATATGGAATTGGAGGAGTACTGGGTGTCCAGTTTCACAAAAGACAGATCTTGTTACATTTG	3480		
QY 3481	CCACAGGCAGTGTCTTTTTTCCGGGGTAACCTTGGGAAAGCTTTTCCGTGAGTTCACTGC	3540		

```

FEATURES
source
  i. .1099
  Location/Qualifiers
    origin quality sequence dbp. 000.
    /organism="Mus musculus"
    /strain="FVB/N"
    /db_xref="taxon:10090"
    /clone="IMAGE:4487204"
    /clone_lib="NCI_CGAP_Maml"
    /tissue_type="tumor, biopsy sample"
    /dev_stage="10 months, virgin"
    /lab_host="DH10B"
    /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Sali;
    Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
    Library constructed by Life Technologies. Investigator
    providing samples: Gilbert Smith, NIH"
  BASE COUNT      370 a      219 c      289 g      221 t
  ORIGIN

Query Match      14.3%; Score 544; DB 12; Length 1099;
Best Local Similarity 88.5%; Pred. No. 7.1e-122;
Matches 680; Conservative 0; Mismatches 80; Indels 8; Gaps 8;

```

[illegible]

QY 2978 AACTAATTAAGAGAGTCGAGAGACTTATCTC-CGGACAAAGAGTCTTAACCCATATTT 3036
Db 61 AAGCTGATTAAAGGAGTTGAAGGCTGATCTCGTGGACAACAAGGTCTTAATCCATATTT 120
QY 3037 AACTTTGAACACCTCTGGTAGTGAACAATCTTATAGATCTGTCTCCTGATGATAAGA 3096
Db 121 AACTCTAAACAACCTCTGGTAGTGAACAATCTCATAGATCTGTCTCCTGATGATAAGA 180
QY 3097 GTTTCAGTCTGTGGAGGAAGAGATGCAAGTACAGATTCGAGAGCACAGAGATGGAGTCA 3156
Db 181 ATTTCAGTCTGTGAAGAAGAGATGCAGAGTACTGTGCGGAGACACAGAGATGGTGTCA 240
QY 3157 TGCAGGTGGAATCTTCAACAGATACAAATATCTCAAGATTCAGAAGGTTTGTAAACAAGAA 3216
Db 241 CGCAGGCGGCTCTTCAACAGATACAAATCTCAAGATTCAGAAGGTTTGTAAACAAGAA 300
QY 3217 ACTATGGGAAGATACACTCACCAGGAGAAAGAGTTTCTGAAGAAACACCAACCATGC 3276
Db 301 ATTGTGGGAAGATATACACACCGGAGAAAGAGTTTCTGAAGAAACACCAACCATGC 360
QY 3277 CAATGAACGAATCTATTTTCATGGTCTCCTTTTGTGAATGCAATATCCACAAGGCTT 3336
Db 361 AAGCAAGAGATGTTATTTTCATGGTCTCCTTTTGTGAATGCGATATCCATAGGCTT 420
QY 3337 TGATGAAGGCATCGGTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGTGAAAA 3396
Db 421 TGATGAAGGCATCGCATACATAGTGGCATGTTATGGAGCTGGAATTTATTTTGTGAGAA 480
QY 3397 CTCTTCCAAAAGCAATATATATATGGAATTTGGAGGAGTACTGGGT-GTCCAGTTC 3455
Db 481 CTCTTCCAAAAGCAATATATATGTTATGGAATCGGAGGTGGCACCAGGATGGTCCCAATC 540
QY 3456 ACAAGACAGATCTTGTACATTTGCCACAGCAGCTGCTTTTGGCGGTAAACCTTGG 3515
Db 541 ACAAGACAGATCGTGTACATTTGTCACAGCAGCTGCT-GTTTGTGAGTAACCTTGG 599
QY 3516 GAAAGCTTTTCCGAGTTCAGTGCATGAAATGGAATGGCACATTCCTCCAGGTCACTAC 3575
Db 600 GAAAGCTTTTCCGAGTTCAGGCAATGAAATGGCACATTCCTCCCGGGCATCACT 659
QY 3576 CAGTCACTGTAGGCCAGTGAATGGCCTAGCATTAGCTGAATATGTTATTACAGAG 3635
Db 660 CGTCACTGGGCGG-CCAGTGTAAATGG-CTAGCATTAG-TGAATATG-TATTACAGAG 715
QY 3636 GAGACAGGCTTATCTGTAGTATTTAATTACTTACCAGATTAAGGC 3683
Db 716 GAAACAGG-TTATCTGAAATCTAATTACTTTACCGATGGTAGGCC 762

Search completed: February 11, 2003, 21:15:20
Job time : 4968.65 secs

REFERENCE	2 (bases 1 to 6028)									
AUTHORS	Chi, N.-W. and Iodish, H. F.									
TITLE	Direct Submission									
JOURNAL	Submitted (27-SEP-2000) Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA									
FEATURES	Location/Qualifiers									
source	1..6028	/organism="Homo sapiens"								
		/db_xref="taxon:9606"								
		/chromosome="10"								
		/map="10q23.1"								
		/note="STS WI-16054"								
gene	1..6028	/gene="TNKS-1"								
	252..3752	/gene="TNKS-2"								
CDS		/gene="TNKS-2"								
		/note="poly(ADP-ribose) polymerase containing ankyrin repeats and SAM (sterile alpha module) motif"								
		/codon_start=1								
		/product="tankyrase-2"								
		/protein_id="AAK82330.1"								
		/db_xref="GI:15042552"								
		/translation="MSGRCAGGACGACASAAAEVAFPAARELFEACRNGDVERVKRLVTPKVNSDPTAGRSKSTPLHFAAGGRKDVVEYLLONGANQARDQGLIPLHNACSFVHAEVNIILRHGADPNARDNWNPTPLHEAAIKGIDCVILLQHGAEPTIRNTDGRTAHLADPESAKAVLGEYKKDLLEFSARGNEEKMWALLTPLNVNCHASDGRKSTPLHLAAGNRYKIVOLLQHGADVHAKDKGLDPLVHNCSYCHYEVTPELLYKHGACVAMDLWQFTPLHEAAKSNVCEVCSLLSYSCADPTLLNCHNKSANDLAPTPOLKERLAYEFKSHSLLQAAREADVTRIKKHLSELMWNFKIPQTHETALHCAASPYPRKQICELLRLKGANINBKTGEFPLHVAEKAHNDVVEVVVHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLSYSCDPNLISLQAGTALQMGNEENVQQLQEGI SLGNSEADROLLEAAKAGDVETVKKICTQSVNCRDIIEGRQSTPLHFAAGGYNRVSYEYLQHGADVHAKDGGGLVPLHNA CSYGHVEABLLYKHGAGVNVADLWKFTPLHEAAAKGYEICKLLQHGADPTKKNRDGNTPDLVDKGDFTDIDLQHDAAALDAAKGCGLARVKLSSPDVNVNCRDTCGRHSTPLHLAAGYNLDEAVYILQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYACVNA TDKWAETPLHEAAKQGTQICALLAHGADPTLKNQEGTPLDLVSADVDVSALLTAMPPSALPSCYKQVINGRPGATADALSPPSGPSSLASLNSLDSFSELSVSVSSSTEGASSLKEKVEGVDPSTFVFNRLGLEHMDIFEREQITLDLVEMGHKELKEIGNAVCHRRKLIKAGVERLSQOGLNPLYLTNTSGSTLIDLSFDDKEFQSEHEEMQTSVNRHSDGHAGGIYFNRYNLIKIQVCNKLWERYTHRRKEVSENNHNHNERMLFHGSPGVNNAIIHKGFDERHAYIGGMPGAGIYFAENSKSNOYVYIGGTCGCPVHKORSYCYTHRQLLPFCVTYLGKSFLOFSAMKMAHSPPGHHSVTGRPSVNGLAELAEVYIRGEQAYPEYLLTYOIMRPEGMVDG"								
		/gene="TNKS-2"								
		/polyA_signal								
BASE COUNT	1724 a	1180 c	1336 g	1788 t						
ORIGIN										
	Query Match 99.4%; Score 3793; DB 9; Length 6028;									
	Best Local Similarity 99.8%; Pred. No. 0;									
	Matches 3807; Conservative 0; Mismatches 5; Indels 1; Gaps 1;									
QY	1	CGCGCTGCTCGCCCGCGGGGCGACGCGGGGGGCGAGGAGCCACGAGGGGCGCGCG 60								
Db	29	CGCGCTGCCTCGCCGCGCGGGGCGACGCGGGGGCGAGGAGCCACGAGGGGCGCGCG 88								
QY	61	TGGCGCGGGCCCATGGGACTGCGCGCGATCGGTTGACAGAGGAGGACGACGCGCGCG 120								
Db	89	TGGCGCGGG - CCATGGGACTGCGCGGATCGGTTGACAGAGGAGGACGACGCGCGCG 147								
QY	121	CCCTGAGCGGCTCTTCTCCGGGGGGCTCGGCCCTCTCTGCTCGCGGGGCGCGGGCTCTCTGC 180								
Db	148	CCCTGAGCGGCTCTTCTCCGGGGGGCTCGGCCCTCTCTGCTCGCGGGGCGCGGGCTCTCTGC 207								
QY	181	TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGGCCAGGATCATGTCGGGGTCCGCGCT 240								
Db	208	TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGGCCAGGATCATGTCGGGGTCCGCGCT 267								
QY	241	GCGCCGCGGGGAGCGCTGCGGAGCGCGCGCGCGGCGGCTGAGCGCGCGCGCGCC 300								
Db	268	GCGCCGCGGGGAGCGGCTTGCBCGAGCGCGCGCGCGGCGGCGGCTGAGCGCGCGCGCC 327								
QY	301	GAGAGCTGTTTCGAGGCGTGC CGCAACGGGACSTGGAACAGTCAAGAGGCTGTGTGACG 360								

Db	328	GAGAGCTGTTTCGAGGCGTCCGCCAACGCGGACGTGGAACAGCTCAAGAGCGTGGTGACGC	387
Qy	361	CTGAGAAGGTGAACAGCGCGACACACGCGCGGAGGAAATCCACCCGCTGCACCTTCGCGC	420
Db	388	CTGAGAAGGTGAACAGCGCGACACACGCGCGGAGGAAATCCACCCGCTGCACCTTCGCGC	447
Qy	421	CAGGTTTTGGCGGAAGAGCGTAGTTGAATATTTCGTTTCAGAAATGGTGCAAATGTCCAAG	480
Db	448	CAGGTTTTGGCGGAAGAGCGTAGTTGAATATTTCGTTTCAGAAATGGTGCAAATGTCCAAG	507
Qy	481	CACGTGATGATGGGGGCTTATTCCTTCATCAATCATGCTTCCTTTGGTGTATGCTGAAG	540
Db	508	CACGTGATGATGGGGGCTTATTCCTTCATCAATGCGATGCTTTTGGTCACTGCTGAAG	567
Qy	541	TAGTCAATCTCCTTTTGGCAGACTGGTGCAGACCCCAATGCTCGAGATAATTGGAAATTATA	600
Db	568	TAGTCAATCTCCTTTTGGCAGACTGGTGCAGACCCCAATGCTCGAGATAATTGGAAATTATA	627
Qy	601	CTCTCTCTCCATGAAGCTGCAATTAAGGGAAGAGATTGATCTTTTGTCATTTGCTGTTTACAGC	660
Db	628	CTCTCTCTCCATGAAGCTGCAATTAAGGGAAGAGATTGATCTTTTGTCATTTGCTGTTTACAGC	687
Qy	661	ATGAGCTGAGCCAAACCATCCGAAATACAGATCGAAGGACAGACATTGGATTAGCAGATC	720
Db	688	ATGAGCTGAGCCAAACCATCCGAAATACAGATCGAAGGACAGACATTGGATTAGCAGATC	747
Qy	721	CATCTGCCAAGCAGTGCTTTACTGGTGAATATAAGAAAGATGAACCTCTTAGAAAGTGCCA	780
Db	748	CATCTGCCAAGCAGTGCTTTACTGGTGAATATAAGAAAGATGAACCTCTTAGAAAGTGCCA	807
Qy	781	GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCAITTAATGTCAACTGCCACG	840
Db	808	GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCAITTAATGTCAACTGCCACG	867
Qy	841	CAAGTGATGCGAAGAGTCAACTCCATTCATTTGGCAGCAGGATATACAGAGTAAAGA	900
Db	868	CAAGTGATGCGAAGAGTCAACTCCATTCATTTGGCAGCAGGATATACAGAGTAAAGA	927
Qy	901	TTGTFACAGCTTTACTGCAACATGGAGCTGATGCTCCATGCTAAAGATAAAGTGATCTGG	960
Db	928	TTGTFACAGCTTTACTGCAACATGGAGCTGATGCTCCATGCTAAAGATAAAGTGATCTGG	987
Qy	961	TACCATTTACACATAGCCTGTTCTTATTTGGTTCATATGAAGTAAGTGAAGTTCGTCACG	1020
Db	988	TACCATTTACACATAGCCTGTTCTTATTTGGTTCATATGAAGTAAGTGAAGTTCGTCACG	1047
Qy	1021	ATGTTGCTCTGTGTAATGCAATGCACTGTGGCAATTCACCTCCTTCATGAGGCAGCTT	1080
Db	1048	ATGTTGCTCTGTGTAATGCAATGCACTGTGGCAATTCACCTCCTTCATGAGGCAGCTT	1107
Qy	1081	CTAAGAACAGGGTTGAAGTATGTTCTCTTCTTAAGTTATGGTCAGACCCCAACACTGC	1140
Db	1108	CTAAGAACAGGGTTGAAGTATGTTCTCTTCTTAAGTTATGGTCAGACCCCAACACTGC	1167
Qy	1141	TCAATTTGTCACATAAAGTGCTATAGACTTGGCTCCACACCCACAGTTTAAAGAAAGAT	1200
Db	1168	TCAATTTGTCACATAAAGTGCTATAGACTTGGCTCCACACCCACAGTTTAAAGAAAGAT	1227
Qy	1201	TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGGAGAAGCTGATGTTACTC	1260
Db	1228	TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGGAGAAGCTGATGTTACTC	1287
Qy	1261	GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAACACATGAAA	1320
Db	1288	GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAACACATGAAA	1347
Qy	1321	CAGCATTTGCATTTGCTGCTGCAATCTCCATATCCCAAAAAGAACAAATATGTGAACCTG	1380
Db	1348	CAGCATTTGCATTTGCTGCTGCAATCTCCATATCCCAAAAAGAACAAATATGTGAACCTG	1407
Qy	1381	TGCTTAAGAAAAAGGACCAACATCAATTGAAAGCACTAAAGAAATTTCTTGACTCCTCTGCAC	1440

Db 1408 TGCTAAGAAAAGGAGCAAAACATCAATGAAAGAGCTAAAGAAATCTTGACTCCTCTGCACG 1467
Qy 1441 TGGCATCTGAGAAAGCTCAATAATCATCTTCTGAAGTAGTGGTGAACATGAAGCAAGG 1500
Db 1468 TGGCATCTGAGAAAGCTCAATAATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGG 1527
Qy 1501 TTAATGCTCTGGATAAATCTTTGGTTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATC 1560
Db 1528 TTAATGCTCTGGATAAATCTTTGGTTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATC 1587
Qy 1561 TACAAACCTGCCGCTACTCTGAGCTATGCGTGTGATCTCAACATTAATCCCTTCAGG 1620
Db 1588 TACAAACCTGCCGCTACTCTGAGCTATGCGTGTGATCTCAACATTAATCCCTTCAGG 1647
Qy 1621 GCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCTCTCAAGAGGGTATCT 1680
Db 1648 GCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCTCTCAAGAGGGTATCT 1707
Qy 1681 CATTAGSTAAATTCAGAGCGACAGACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCG 1740
Db 1708 CATTAGSTAAATTCAGAGCGACAGACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCG 1767
Qy 1741 AAACCTGTAAGAAACTGTGTACTCTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTC 1800
Db 1768 AAACCTGTAAGAAACTGTGTACTCTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTC 1827
Qy 1801 AGTCTACACCACTTCAATTTGACGCTGGGTATAACAGAGTGTCCGTTGGTGAATATCTGC 1860
Db 1828 AGTCTACACCACTTCAATTTGACGCTGGGTATAACAGAGTGTCCGTTGGTGAATATCTGC 1887
Qy 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGCGCTTGTAACCTTTGCACAATG 1920
Db 1888 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGCGCTTGTAACCTTTGCACAATG 1947
Qy 1921 CATGCTTTATGGACATTAAGAGTTCGCAACTCTCTTTGAACATGGAGCAGTAGTTA 1980
Db 1948 CATGCTTTATGGACATTAAGAGTTCGCAACTCTCTTTGAACATGGAGCAGTAGTTA 2007
Qy 1981 ATGAGCTGATTTATGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATG 2040
Db 2008 ATGAGCTGATTTATGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATG 2067
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAAGAGGATGGAA 2100
Db 2068 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAAGAGGATGGAA 2127
Qy 2101 ATACTCCTTTGCGATCTGTTAAAGATGSAGATACAGATATTCAGATCTGCTTAGGGGAG 2160
Db 2128 ATACTCCTTTGCGATCTGTTAAAGATGSAGATACAGATATTCAGATCTGCTTAGGGGAG 2187
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAGGCTTCTTTAGCCAGAGTGAAGAAGTTGCTTT 2220
Db 2188 ATGCAGCTTTGCTAGATGCTGCCAAGAGGCTTCTTTAGCCAGAGTGAAGAAGTTGCTTT 2247
Qy 2221 CTCCTGATAATGTAATTTGCGCGATACCCCAAGCGACATTTCAACACCTTTACATTTAG 2280
Db 2248 CTCCTGATAATGTAATTTGCGCGATACCCCAAGCGACATTTCAACACCTTTACATTTAG 2307
Qy 2281 CAGCTGGTTATAAATTTAGAGTTGCAGAGTATTTGTTACACACGGAGCTGATCTGA 2340
Db 2308 CAGCTGGTTATAAATTTAGAGTTGCAGAGTATTTGTTACACACGGAGCTGATCTGA 2367
Qy 2341 ATGCCCAAGACAAAGGAGCTTATCTCTTACATAATGACGATCTTACGGGCGATCTAG 2400
Db 2368 ATGCCCAAGACAAAGGAGCTTATCTCTTACATAATGACGATCTTACGGGCGATCTAG 2427
Qy 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGATGTGTCAATGCCACGCAAAATGGGCTT 2460
Db 2428 ATGTAGCAGCTCTACTAATAAAGTATAATGATGTGTCAATGCCACGCAAAATGGGCTT 2487
Qy 2461 TCACACCTTTGACCAAGCAGCCCAAGGAGGACGACACAGCTTTGTGCTTTGTGCTAG 2520
Db 2488 TCACACCTTTGACCAAGCAGCCCAAGGAGGACGACACAGCTTTGTGCTTTGTGCTAG 2547

Qy 2521 CCCATGGAGCTGACCCGACTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db 2548 CCCATGGAGCTGACCCGACTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2607
Qy 2581 CAGCGGATGATGTACGCGCTCTTCTGACAGCAGCAGCCATGCCCATCTGCTGCGCCTCTT 2640
Db 2608 CAGCGGATGATGTACGCGCTCTTCTGACAGCAGCAGCCATGCCCATCTGCTGCGCCTCTT 2667
Qy 2641 GTTACAAAGCTCAAGTGTCTCAATGTTGTGAGAAAGCCAGGAGCCACTGCAGATGCTCTCT 2700
Db 2668 GTTACAAAGCTCAAGTGTCTCAATGTTGTGAGAAAGCCAGGAGCCACTGCAGATGCTCTCT 2727
Qy 2701 CTTTCCAGTCCATCTAGCCCATCAAGCTTTCTGACGCGAGCAGTCTTTGACAACTTTATCTG 2760
Db 2728 CTTTCCAGTCCATCTAGCCCATCAAGCTTTCTGACGCGAGCAGTCTTTGACAACTTTATCTG 2787
Qy 2761 GGAGTTTTTTCAGAACTGCTTCAGTAGTTAGTTTCAAGTGGAAACAGAGGCTGCTCCAGTT 2820
Db 2788 GGAGTTTTTTCAGAACTGCTTCAGTAGTTAGTTTCAAGTGGAAACAGAGGCTGCTCCAGTT 2847
Qy 2821 TGGAGAAAAGAGGCTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGAAATCTTG 2880
Db 2848 TGGAGAAAAGAGGCTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGAAATCTTG 2907
Qy 2881 GACTTGAGCACCTTAATGGATATATTTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db 2908 GACTTGAGCACCTTAATGGATATATTTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2967
Qy 2941 AGATGGGCGACAGAGAGCTGAAGGAGTGGAAATCAATGCTTATGGACATAGGCGACAAC 3000
Db 2968 AGATGGGCGACAGAGAGCTGAAGGAGTGGAAATCAATGCTTATGGACATAGGCGACAAC 3027
Qy 3001 TAATTAAGSAGTGCAGAGACTTATCTCCGACAAACAGTCTTAACCCATATTTAACTT 3060
Db 3028 TAATTAAGSAGTGCAGAGACTTATCTCCGACAAACAGTCTTAACCCATATTTAACTT 3087
Qy 3061 TGAACACCTCTGCTAGTGGAAACAAATTTATAGATCTGCTCTCTGATGAAGAAGTTTC 3120
Db 3088 TGAACACCTCTGCTAGTGGAAACAAATTTATAGATCTGCTCTCTGATGAAGAAGTTTC 3147
Qy 3121 AGTCTGTGGAGAGAGATGCAAAAGTACAGTTTCGAGAGCAGAGATGAGGTCATGAG 3180
Db 3148 AGTCTGTGGAGAGAGATGCAAAAGTACAGTTTCGAGAGCAGAGATGAGGTCATGAG 3207
Qy 3181 GTGCAATCTTCAACAGATACAATATTCTCAAGATTCAAGAGTTTGTAAACAAGAACTAT 3240
Db 3208 GTGCAATCTTCAACAGATACAATATTCTCAAGATTCAAGAGTTTGTAAACAAGAACTAT 3267
Qy 3241 GGGAAAGATACACTCACCGGAGAAAAAGAGTTTCTGAAGAAAAACCAACCATGCCAATG 3300
Db 3268 GGGAAAGATACACTCACCGGAGAAAAAGAGTTTCTGAAGAAAAACCAACCATGCCAATG 3327
Qy 3301 AACGAATGCTATTTCATGGGTCTCTCTTTGTGAATGCAATATCCACAAGGCTTTGATG 3360
Db 3328 AACGAATGCTATTTCATGGGTCTCTCTTTGTGAATGCAATATCCACAAGGCTTTGATG 3387
Qy 3361 AAAGGATGCTACATAGGTGATGTTTGGAGCTGCGCATTTATTTTGTGTAAGAACTCTT 3420
Db 3388 AAAGGATGCTACATAGGTGATGTTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTT 3447
Qy 3421 CCAAAAGCAATCAATATGATATGGAATTTGGAGAGGTACTGGGTCTCCAGTTCCAAAG 3480
Db 3448 CCAAAAGCAATCAATATGATATGGAATTTGGAGAGGTACTGGGTCTCCAGTTCCAAAG 3507
Qy 3481 ACAGATCTTGTACATTTGCCAGAGCAGTGTCTTTTGGCGGGTAACTTTGGGAAAGT 3540
Db 3508 ACAGATCTTGTACATTTGCCAGAGCAGTGTCTTTTGGCGGGTAACTTTGGGAAAGT 3567
Qy 3541 CTTTCTGTCAGTTTCAGTGGCAATGAAATGACATCTCTCTCCAGGTCACTCACTAGTCA 3600
Db 3568 CTTTCTGTCAGTTTCAGTGGCAATGAAATGACATCTCTCTCCAGGTCACTCACTAGTCA 3627

QY 3721 GATAAATAGTTATTTTAAAGAACTAATTCCTACTGAACCTAAAAATCATCAAAAGCAGCAGTG 3780
|||||
Db 3748 GATAAATAGTTATTTTAAAGAACTAATTCCTACTGAACCTAAAAATCATCAAAAGCAGCAGTG 3807
|||||
QY 3781 GCCTACAGTTTCTCTTCTGCTGAAAAAAA 3813
|||||
Db 3808 GCCTACAGTTTCTCTTCTGCTGAAAAAAA 3840
|||||

RESULT 3
AX062247
LOCUS AX062247 4406 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 106 from Patent WO0100849.
ACCESSION AX062247
VERSION AX062247.1 GI:12540148
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4406)
AUTHORS Christenson, E., Demaggio, A. J., Goldman, P. S. and Mcelligott, D. L.
TITLE Tankyrase2 materials and methods
JOURNAL Patent: WO 0100849-A 106 04-JAN-2001;
ICOS CORPORATION (US)

FEATURES
Location/Qualifiers
1..4406
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS
1..3790
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC25062.1"
/db_xref="GI:12540149"

/translation="ELAGALPASAAAFODPDGFGALPPPPRSGRSGAGSPARGARG
GHGTPDVTAGSQARALSSPGGLALLAGPGLLRLLALLAVAAADMSGRRC
AGGAACASAAAEVPEAPARELFECRNGDVERKRLVTPKEVNSRDTRGRKSTPLHF
AAGFRKDVYELQNGANVARDDGLIPLHNACSPGHAEVYNLLRHGADPNARDN
WNYPLHEAAIKGKTIDVCTVLIQHGAEPTIRTDGRTALDLADPSAKAVLTGEYKDE
LLESARNEEMKALLPLPNVCHASDGRKSTPLHLAAGNRVKIIVOLLQHGADVH
AKDGDVPLHNACSYGHEVTELLVHKACVANDLHQFTPLHEAASKNRVEVCSLL
LSYGADPILLNHNKSAIDLAPTQLKERLAYEFKGLSLQARADVTRIKKLSLE
MYNKHPTHTALHCAASAPYKPKQICELLLRKGANINEKTRFELPLHVASDKAH
NDVEVYVYKHEAKNALNDGQTSILHRAAYCGLQTCRLLSYGDNPNIISLQFTAL
QMGNEVQQLQEGTSLGNSEADRLLEAAKAGDVEVTKICTVQSVNCRDIEGRQST
PLHFAAGYNRVSVVYLLQHGADVHAKDGLVPLHNACSYGHYEVALLVYKHGAVN
VADLWFTPLHEAAKGYEICKLLOHGADPTKKNRDGNTPDLVKGDDTDIQDLLR
GDAALLDAKKGLCLARYKLSLPDNNVNRDQGRHSTPLHLAAGNNLEAVEYLLQHG
ADVNAQDGGLLPLHNAASYGHVDAALLIKYNACVNATDKWAFPLPHEAAKGRQL
CALLLAHGADPTLKNEQGTPLDLSADDSALITAMPSPALPSCYQVILVYRSP
GATADALSPPSSLSAASDLNLSGFSSELSVSSSGTEGASLLEKKEVPQVDF
SITQVRNLGLEHMDIFERQGITLDVLVEMGHKELKEIGINAYGHRKLLKGVRLI
SGOGLNPLTLNTSGCTILLDPSDDKRFQSVVEEMQSTVYRHRDGHAGGIFNRY
NLLKQVKNKLMERYTHRRKEVSEENHNRMFLHGFSPFNAILHKGFDERHAY
IGMKFGAGIYFAENSNSQNYVYGGGTGCPVHKDRSCYICHROLLEFCVTLRKSFL
QFSAMKMAHSPPHSHSVTGRPSVNGLAIAEYVYRGEQAYPEYLLITYQIMRPEGMVDG
"

BASE COUNT 1268 a 910 c 1054 g 1174 t
ORIGIN
Query Match 99.4%; Score 3791.4; DB 6; Length 4406;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 CGCGTGTCCCGCGCGGCGAGCGGGGCGAGGAGCCCGAGCGGGCGCGG 60
|||||
Db 67 CGCGTGTCCCGCGCGGCGAGCGGGGCGAGGAGCCCGAGCGGGCGCGG 126
|||||
QY 61 TGGGCGGCGCCATGGGACTGCGCGCGATCCCGGTGCACACAGGAGCCCAAGCGCGCGG 120
|||||
Db 127 TGGGCGCGG-CCATGGGACTGCGCGCGATCCCGGTGCACACAGGAGCCCAAGCGCGCGG 185
|||||

QY 121 CCTGAGCGCGTCTTCTCGGGGGGCGCTCGCCCTCTCTGCTCGCGGGCGCGGGCTCTCTGC 180
|||||
Db 186 CCTGAGCGCGTCTTCTCGGGGGGCGCTCGCCCTCTCTGCTCGCGGGCGCGGGCTCTCTGC 245
|||||
QY 181 TCCGGTTGCTGGCGCTGTGTGCTGGTGTGGCGGCGCAGGATCATGTTCGGGTCTCGCGCT 240
|||||
Db 246 TCCGGTTGCTGGCGCTGTGTGCTGGTGTGGCGGCGCAGGATCATGTTCGGGTCTCGCGCT 305
|||||
QY 241 GCGCGCGGGGAGCGCGCTCGCGGAGCGCGCGCGGCGGCGGCTGGAGCGCGCGCGCC 300
|||||
Db 306 GCGCGCGGGGAGCGCGCTCGCGGAGCGCGCGCGGCGGCGGCTGGAGCGCGCGCGCC 365
|||||
QY 301 GAGAGCTGTTGAGAGCGTGGCCCAACGGGACGCTGGAACAGTCAAGAGGCTGCTGACGC 360
|||||
Db 366 GAGAGCTGTTGAGAGCGTGGCCCAACGGGACGCTGGAACAGTCAAGAGGCTGCTGACGC 425
|||||
QY 361 CTGAGAAGGTGAACAGCGCGGACCGGGGAGGAAATCCACCCCGCTGCACTTCGCGC 420
|||||
Db 426 CTGAGAAGGTGAACAGCGCGGACCGGGGAGGAAATCCACCCCGCTGCACTTCGCGC 485
|||||
QY 421 CAGGTTTTCGGCGGAAGACGTAGTTGAATATTTGCTTCAGAATGCTGCAAAATGTCCAAG 480
|||||
Db 486 CAGGTTTTCGGCGGAAGACGTAGTTGAATATTTGCTTCAGAATGCTGCAAAATGTCCAAG 545
|||||
QY 481 CACGTGATGATGGGGCCTTATTCCTCTTCAATATGCATGCTCTTTTGGTCTATGCTGAAG 540
|||||
Db 546 CACGTGATGATGGGGCCTTATTCCTCTTCAATATGCATGCTCTTTTGGTCTATGCTGAAG 605
|||||
QY 541 TAGTCAATCTCTTTTGCACATGTTGCGAGACCCCAATGCTCGAGATAATTGGAATATA 600
|||||
Db 606 TAGTCAATCTCTTTTGCACATGTTGCGAGACCCCAATGCTCGAGATAATTGGAATATA 665
|||||
QY 601 CTCCTCTCCATGAAGCTGCAATTAAGAGAAAGATTGATGTTTTCATTTGCTCTTACAGC 660
|||||
Db 666 CTCCTCTCCATGAAGCTGCAATTAAGAGAAAGATTGATGTTTTCATTTGCTCTTACAGC 725
|||||
QY 661 ATGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGAGACAGCATTTGGATTTAGCAGATC 720
|||||
Db 726 ATGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGAGACAGCATTTGGATTTAGCAGATC 785
|||||
QY 721 CATCTGCCAAACGAGTCTACTGCTGAATATGAAGAGATGAACCTCTTAGAAGTGCCA 780
|||||
Db 786 CATCTGCCAAACGAGTCTACTGCTGAATATGAAGAGATGAACCTCTTAGAAGTGCCA 845
|||||
QY 781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACATTAAATGCACTGCCACG 840
|||||
Db 846 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACATTAAATGCACTGCCACG 905
|||||
QY 841 CAAAGTGGCAGAAAGTCACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
|||||
Db 906 CAAAGTGGCAGAAAGTCACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 965
|||||
QY 901 TTGTACAGCTGTACTGCAACATGGAGCTGATGCTCATGCTAAAGATAAAGGTGATCTGG 960
|||||
Db 966 TTGTACAGCTGTACTGCAACATGGAGCTGATGCTCATGCTAAAGATAAAGGTGATCTGG 1025
|||||
QY 961 TACCATTACACAATGCTGTTCTTATGCTCATTAAGTAACTGAACCTTTTGGTCAAGC 1020
|||||
Db 1026 TACCATTACACAATGCTGTTCTTATGCTCATTAAGTAACTGAACCTTTTGGTCAAGC 1085
|||||
QY 1021 ATGGTGCTGTGTAATGCAATGGACTTTGGCAATTCTACTCTTTCATGAGGAGCTT 1080
|||||
Db 1086 ATGGTGCTGTGTAATGCAATGGACTTTGGCAATTCTACTCTTTCATGAGGAGCTT 1145
|||||
QY 1081 CTAAGAACAGGTTGAAGTATGTTCTTCTTAAGTTATGCTGAGACCCCAACTGC 1140
|||||
Db 1146 CTAAGAACAGGTTGAAGTATGTTCTTCTTAAGTTATGCTGAGACCCCAACTGC 1205
|||||
QY 1141 TCAATTTGCACAAATAAAGTGTATAGACTTGGCTCCACACACACAGTTTAAAGAAAGAT 1200
|||||
Db 1206 TCAATTTGCACAAATAAAGTGTATAGACTTGGCTCCACACACACAGTTTAAAGAAAGAT 1265
|||||
QY 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTC 1260
|||||

QY 961 TACCATTACACAAATGCCCTGTTCTTATGGTCATATGAAGTAAGTAACTGAACCTTTTGGTCAAGC 1020
Db 1612 TACCATTACACAAATGCCCTGTTCTTATGGTCATATGAAGTAAGTAACTGAACCTTTTGGTCAAGC 1671
QY 1021 ATGGTGCCGTGTGAATGAATGCAATGACCTGTGGCAATTCACCTCTCTTTCATGAGGACGCTT 1080
Db 1672 ATGGTGCCGTGTGAATGCAATGACCTGTGGCAATTCACCTCTCTTTCATGAGGACGCTT 1731
QY 1081 CTAGAACAGGGTTGAAGTATGTTCTCTTCTTAAAGTTATGGTGCAGACCCACACACTGC 1140
Db 1732 CTAGAACAGGGTTGAAGTATGTTCTCTTCTTAAAGTTATGGTGCAGACCCACACACTGC 1791
QY 1141 TCAATTGTCAACAATAAAGTCTATAGACTTGGCTGCCACACACAGCTTAAAGAAAGAT 1200
Db 1792 TCAATTGTCAACAATAAAGTCTATAGACTTGGCTGCCACACACAGCTTAAAGAAAGAT 1851
QY 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGGAAGCTGATGTTACTC 1260
Db 1852 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGGAAGCTGATGTTACTC 1911
QY 1261 GAATCAAAAAACATCTCTCTCGAATGGTGAATTTCAAGCATCTCTCAACACATGAAT 1320
Db 1912 GAATCAAAAAACATCTCTCTCGAATGGTGAATTTCAAGCATCTCTCAACACATGAAT 1971
QY 1321 CAGCATTTGCTGCTGCTGCTCATCTCCATATCCCAAGAAAGCAAAATATGTGAAGTGT 1380
Db 1972 CAGCATTTGCTGCTGCTGCTCATCTCCATATCCCAAGAAAGCAAAATATGTGAAGTGT 2031
QY 1381 TGCTAAGAAAAGGAGCAACATCAATGAAAGACTTAAAGAAATCTTGACTCCTCTGCAAG 1440
Db 2032 TGCTAAGAAAAGGAGCAACATCAATGAAAGACTTAAAGAAATCTTGACTCCTCTGCAAG 2091
QY 1441 TGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTASTGGTGAACATGAAGCAAGG 1500
Db 2092 TGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTASTGGTGAACATGAAGCAAGG 2151
QY 1501 TTAATGCTCTGATATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATC 1560
Db 2152 TTAATGCTCTGATATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATC 2211
QY 1561 TACAACCTTGGCCCTACTCTCGAGCTATGGGTGTGATCCTTAACATATATCCCTTCAGG 1620
Db 2212 TACAACCTTGGCCCTACTCTCGAGCTATGGGTGTGATCCTTAACATATATCCCTTCAGG 2271
QY 1621 GCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCTCCAGAGGGTATCT 1680
Db 2272 GCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCTCCAGAGGGTATCT 2331
QY 1681 CATTTAGTAAATTCAGAGGCAGACACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGTG 1740
Db 2332 CATTTAGTAAATTCAGAGGCAGACACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGTG 2391
QY 1741 AAATGTAAAAAATCTGTACTGTTTACAGTGTCAACTGCAGACACATTTGAAGGGCGTC 1800
Db 2392 AAATGTAAAAAATCTGTACTGTTTACAGTGTCAACTGCAGACACATTTGAAGGGCGTC 2451
QY 1801 AGTCTACACCTTCAATTTTGCAGCTGGGTATACAGAGTGTCCGTGGGAATATCTGC 1860
Db 2452 AGTCTACACCTTCAATTTTGCAGCTGGGTATACAGAGTGTCCGTGGGAATATCTGC 2511
QY 1861 TACAGCATGGAGCTGATGTGCATGTAAAGATAAAGAGGCGCTTGTACCTTTGCAACAATG 1920
Db 2512 TACAGCATGGAGCTGATGTGCATGTAAAGATAAAGAGGCGCTTGTACCTTTGCAACAATG 2571
QY 1921 CATGTTCTTATGGCAATATGAAGTTGCAAGCTTCTGTTTAAACATGGAGCAGTAGTTA 1980
Db 2572 CATGTTCTTATGGCAATATGAAGTTGCAAGCTTCTGTTTAAACATGGAGCAGTAGTTA 2631
QY 1981 ATGTAGCTGATTTATGGAATTTACACCTTTACATGAACAGCAGCAAAAGAAATATG 2040
Db 2632 ATGTAGCTGATTTATGGAATTTACACCTTTACATGAACAGCAGCAAAAGAAATATG 2691

QY 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTCAGACCTTACCAAAAAAACACAGGATGGAA 2100
Db 2692 AAATTTGCAAACTTCTGCTCCAGCATGGTCAGACCTTACCAAAAAAACACAGGATGGAA 2751
QY 2101 ATACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAG 2160
Db 2752 ATACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAG 2811
QY 2161 ATGACGCTTTGCTAGATGCTGCCAAAGAGGTTGTTTACCGAGAGTGAAGAGTTCTCTT 2220
Db 2812 ATGACGCTTTGCTAGATGCTGCCAAAGAGGTTGTTTACCGAGAGTGAAGAGTTCTCTT 2871
QY 2221 CTCTGATATATGTAATTTGCCGATACCCAGGACAGATTCACACCTTTACATTTAG 2280
Db 2872 CTCTGATATATGTAATTTGCCGATACCCAGGACAGATTCACACCTTTACATTTAG 2931
QY 2281 CAGCTGGTTATATATATTTAGAAGTTGCAGAGTATTTGTTTACAACACGAGCTGATGTA 2340
Db 2932 CAGCTGGTTATATATATTTAGAAGTTGCAGAGTATTTGTTTACAACACGAGCTGATGTA 2991
QY 2341 ATGCCAAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db 2992 ATGCCAAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 3051
QY 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGATGTGTCAATGCCAGGACAAATGGGCTT 2460
Db 3052 ATGTAGCAGCTCTACTAATAAAGTATAATGATGTGTCAATGCCAGGACAAATGGGCTT 3111
QY 2461 TCACACCTTTGCAGGAGAGAGCCCAAGGAGGACACACAGCTTTGCTTCTGCTGCTAG 2520
Db 3112 TCACACCTTTGCAGGAGAGAGCCCAAGGAGGAGACACAGCTTTGCTTCTGCTGCTAG 3171
QY 2521 CCATGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db 3172 CCATGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 3231
QY 2581 CAGCGGATGATGTCAGGCTCTTCTTGACAGAGCCATGCCCCCATGCTCTGCCCTCTT 2640
Db 3232 CAGCGGATGATGTCAGGCTCTTCTTGACAGAGCCATGCCCCCATGCTCTGCCCTCTT 3291
QY 2641 GTTACAGCTCAAGTCTCAATGGTGTGAGAGCCAGGAGCCAGGACCTGACAGATGCTCTCT 2700
Db 3292 GTTACAGCTCAAGTCTCAATGGTGTGAGAGCCAGGAGCCAGGACCTGACAGATGCTCTCT 3351
QY 2701 CTTGAGTCTCATCTAGCCCCATCAAGCCCTTCTGCGAGCCAGCAGCTTTGACAACCTTATCTG 2760
Db 3352 CTTGAGTCTCATCTAGCCCCATCAAGCCCTTCTGCGAGCCAGCAGCTTTGACAACCTTATCTG 3411
QY 2761 GGAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAAACAGAGGGTCTTCCAGTT 2820
Db 3412 GGAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAAACAGAGGGTCTTCCAGTT 3471
QY 2821 TGGAGAAAAGAGGAGTTCCAGGAGTAGATTTTACATAACTCAATTCCTAAGGAATCTTG 2880
Db 3472 TGGAGAAAAGAGGAGTTCCAGGAGTAGATTTTACATAACTCAATTCCTAAGGAATCTTG 3531
QY 2881 GACITGAGCAGCTAATGGATATATTTGAGAGAGAACAGATCAGCTTTGGATGTATTAAGTTG 2940
Db 3532 GACITGAGCAGCTAATGGATATATTTGAGAGAGAACAGATCAGCTTTGGATGTATTAAGTTG 3591
QY 2941 AGATGGGACACAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGACACAAAC 3000
Db 3592 AGATGGGACACAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGACACAAAC 3651
QY 3001 TAATTAAGGAGCTCGAGAGACTTATCTCCGACACAAAGGCTTTAACCCATATTTAACTT 3060
Db 3652 TAATTAAGGAGCTCGAGAGACTTATCTCCGACACAAAGGCTTTAACCCATATTTAACTT 3711
QY 3061 TGAACACCTCTGGTAGTGAACAATTTCTTATAGATCTGTCTCTCTGATGATAAAGAGTTTC 3120
Db 3712 TGAACACCTCTGGTAGTGAACAATTTCTTATAGATCTGTCTCTCTGATGATAAAGAGTTTC 3771
QY 3121 AGTCTGTGGAGAAAGAGATGCAAAAGTACAGTTTCGAGAGCAGAGATGGAGGTCATGACG 3180

Db 1262 CTCCTCTCCATGAAGCTCAATTAAAGAAAGATTGATGTTTGCTGCTGTACAGC 1371
QY 661 ATGGAGCTGAGCCAAACCAATCCGAATACAGATGGAAGACAGCATTTGGATTACGACATC 720
Db 1322 ATGGAGCTGAGCCAAACCAATCCGAATACAGATGGAAGACAGCATTTGGATTACGACATC 1381
QY 721 CATCTGCCAAAGCAGTCTACTGGTGAATATAAGAAAGATGAACCTTTAGAAAGTCCCA 780
Db 1382 CATCTGCCAAAGCAGTCTACTGGTGAATATAAGAAAGATGAACCTTTAGAAAGTCCCA 1441
QY 781 GGAGTGGCAATGAAGAAATATGATGGCTTACTACACACCATTAATATGCTCAACCTGCCACG 840
Db 1442 GGAGTGGCAATGAAGAAATATGATGGCTTACTACACACCATTAATATGCTCAACCTGCCACG 1501
QY 841 CAAATGATGGCAAGAAATCACTCAATTAATTTGGCAGCAGGATATAACAGAGTAAGA 900
Db 1502 CAAATGATGGCAAGAAATCACTCAATTTGGCAGCAGGATATAACAGAGTAAGA 1561
QY 901 TTGTACAGCTGTTACTGCAACATGAGCTGATGTCCATGTCTAAAGATAAAGGTGATCTGG 960
Db 1562 TTGTACAGCTGTTACTGCAACATGAGCTGATGTCCATGTCTAAAGATAAAGGTGATCTGG 1621
QY 961 TACCATTACAAATGCCCTGTTCTTATGGTCATTTATGAAGTAACCTTTTGGTCAAGC 1020
Db 1622 TACCATTACAAATGCCCTGTTCTTATGGTCATTTATGAAGTAACCTTTTGGTCAAGC 1681
QY 1021 ATGGTCCCTGTGTAATGCAATGACATGACATTTGGCAGATTCACCTCTTCATCAGGACGCTT 1080
Db 1682 ATGGTCCCTGTGTAATGCAATGACATTTGGCAGATTCACCTCTTCATCAGGACGCTT 1741
QY 1081 CTAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAAAGTTATGGTGCAGACCCACACTGC 1140
Db 1742 CTAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAAAGTTATGGTGCAGACCCACACTGC 1801
QY 1141 TCAATTTGCACAAATAAAGTCTATAGACTTTGGCTCCACACCACAGTTAAAGAAAGAT 1200
Db 1802 TCAATTTGCACAAATAAAGTCTATAGACTTTGGCTCCACACCACAGTTAAAGAAAGAT 1861
QY 1201 TAGCATATGAATTTAAAGCCCACTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTC 1260
Db 1862 TAGCATATGAATTTAAAGCCCACTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTC 1921
QY 1261 GAATCAAAAACATCTCTCTCGAAATGGTGAATTTCAAGCATCCCTCAACACATGAAA 1320
Db 1922 GAATCAAAAACATCTCTCTCGAAATGGTGAATTTCAAGCATCCCTCAACACATGAAA 1981
QY 1321 CAGCATTCGATTTGCTGCTGCATCTCCATATCCCAAAAGAACAAATATGTGAACCTGT 1380
Db 1982 CAGCATTCGATTTGCTGCTGCATCTCCATATCCCAAAAGAACAAATATGTGAACCTGT 2041
QY 1381 TGCCTAAGAAAGGAGCAACATCAATGAAAGACTAAAGAAATTTCTGACTCTCTGCACG 1440
Db 2042 TGCCTAAGAAAGGAGCAACATCAATGAAAGACTAAAGAAATTTCTGACTCTCTGCACG 2101
QY 1441 TGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTGGTGAACATCAAGCAAGG 1500
Db 2102 TGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTGGTGAACATCAAGCAAGG 2161
QY 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATTAATTTGGTGCATC 1560
Db 2162 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATTAATTTGGTGCATC 2221
QY 1561 TACAACCTCGCCCTACTCTGAGCTATGGTGTGATCTTAAACATTAATATCCCTTCAGG 1620
Db 2222 TACAACCTCGCCCTACTCTGAGCTATGGTGTGATCTTAAACATTAATATCCCTTCAGG 2281
QY 1621 GCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCTCCCAAGGGTATCT 1680
Db 2282 GCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCTCCCAAGGGTATCT 2341
QY 1681 CATTAGGTAATTCAGAGGCACAGACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGTGG 1740

Db 2342 CATTAGGTAATTCAGAGGCAGACAGACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGTGG 2401
QY 1741 AAAGTGTAAAAAATCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGCGCTC 1800
Db 2402 AAAGTGTAAAAAATCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGCGCTC 2461
QY 1801 AGTCTACACCACCTTCATTTTGGCAGCTGGGTATTAACAGAGTGTCCGTGGGAATATCTGC 1860
Db 2462 AGTCTACACCACCTTCATTTTGGCAGCTGGGTATTAACAGAGTGTCCGTGGGAATATCTGC 2521
QY 1861 TACAGCATGGAGCTGATGTGCATGTCTAAAGATAAAGAGGCGCTTTGTACCTTTGCACAATG 1920
Db 2522 TACAGCATGGAGCTGATGTGCATGTCTAAAGATAAAGAGGCGCTTTGTACCTTTGCACAATG 2581
QY 1921 CATGTTCTTATGACATTAATGAAGTTGCAGAACTTTCTTGTAAACATGAGCAGTAGTTA 1980
Db 2582 CATGTTCTTATGACATTAATGAAGTTGCAGAACTTTCTTGTAAACATGAGCAGTAGTTA 2641
QY 1981 ATCTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATG 2040
Db 2642 ATCTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATG 2701
QY 2041 AAATTTGCAAACTTCCTCCAGCATGGTGCAGACCTTACCAAAAAACACAGGATGGAA 2100
Db 2702 AAATTTGCAAACTTCCTCCAGCATGGTGCAGACCTTACCAAAAAACACAGGATGGAA 2761
QY 2101 ATACTCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGGAG 2160
Db 2762 ATACTCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGGAG 2821
QY 2161 ATCAGCTTTTCTAGATGCTGCCAAAGGGTGTGTTTAGCCAGAGTGAAGAGTGTCTTT 2220
Db 2822 ATCAGCTTTTCTAGATGCTGCCAAAGGGTGTGTTTAGCCAGAGTGAAGAGTGTCTTT 2881
QY 2221 CTCTGTAAATGTAATTTGCGCGATACCCAGGACAGATTCACACCTTTACATTTAG 2280
Db 2882 CTCTGTAAATGTAATTTGCGCGATACCCAGGACAGATTCACACCTTTACATTTAG 2941
QY 2281 CAGCTGTTAATAATTTAGAAAGTTCAGAGTATTTGTTTACAAACAGGAGCTGATGTA 2340
Db 2942 CAGCTGTTAATAATTTAGAAAGTTCAGAGTATTTGTTTACAAACAGGAGCTGATGTA 3001
QY 2341 ATGCCAAGCAAAAGGAGGACTTATCTTTTACATATGAGCATCTTACGGCATGTAG 2400
Db 3002 ATGCCAAGCAAAAGGAGGACTTATCTTTTACATATGAGCATCTTACGGCATGTAG 3061
QY 2401 ATGTAGCAGCTCTACTAATTAAGTATTAATGCATGTGTCAATGCCAGGACAAATGGGCTT 2460
Db 3062 ATGTAGCAGCTCTACTAATTAAGTATTAATGCATGTGTCAATGCCAGGACAAATGGGCTT 3121
QY 2461 TCAACCTTTGCAGAGCAGCCCCAAAGGAGCAGACACAGCTTTGTGCTTTCTGCTAG 2520
Db 3122 TCAACCTTTGCAGAGCAGCCCCAAAGGAGCAGACACAGCTTTGTGCTTTGCTGCTAG 3181
QY 2521 CCCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db 3182 CCCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 3241
QY 2581 CAGGGATGATGTACGGCTCTTCTGACACAGCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db 3242 CAGGATGATGTACGGCTCTTCTGACACAGCATGCCCCCATCTGCTCTGCCCTCTT 3301
QY 2641 GTTACAAGCCTCAAGTCTCAATGGTGTGAGAAAGGAGCAGGACCTGCGAGATGCTCTCT 2700
Db 3302 GTTACAAGCCTCAAGTCTCAATGGTGTGAGAAAGGAGCAGGACCTGCGAGATGCTCTCT 3361
QY 2701 CTTCAAGTCTCATCTAGCCCATCAAGCTTTCTGAGCCAGCAGCTTGTGACAACTTATCTG 2760
Db 3362 CTTCAAGTCTCATCTAGCCCATCAAGCTTTCTGAGCCAGCAGCTTGTGACAACTTATCTG 3421
QY 2761 GGAGTTTTCAGAACTCTCTTACAGTATGTTAGTTCAAGTGAAGCAGAGGCTGCTCCAGTT 2820
Db 3422 GGAGTTTTCAGAACTCTCTTACAGTATGTTAGTTCAAGTGAAGCAGAGGCTGCTCCAGTT 3481

QY 2821 TGAGAAAAAGGAGTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db 3482 TGCAGAAAAAGGAGTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 3541
QY 2881 GACTTCAGCACCTTAATGGATATATTTTCAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db 3542 GACTTCAGCACCTTAATGGATATATTTTCAGAGAGAACAGATCACTTTGGATGTATTAGTTG 3601
QY 2941 AGATGGGCAACAGGAGCTGAAGCAGATGGAATCAATGCTTATGACATAGGCACAAAC 3000
Db 3602 AGATGGGCAACAGGAGCTGAAGCAGATGGAATCAATGCTTATGACATAGGCACAAAC 3661
QY 3001 TAATTAAGGAGTCGAGAGACTTATCTCCGACAAAGGCTTTAACCCCATATTTAACTT 3060
Db 3662 TAATTAAGGAGTCGAGAGACTTATCTCCGACAAAGGCTTTAACCCCATATTTAACTT 3721
QY 3061 TGAACACCTCTGTAGTGAACAAATTTCTATAGATCTGTCTCTGTATGAATAAGAGTTTC 3120
Db 3722 TGAACACCTCTGTAGTGAACAAATTTCTATAGATCTGTCTCTGTATGAATAAGAGTTTC 3781
QY 3121 AGTCTGTGGAGGAAGAGATGCAAAATACAGTTTCGAGAGCACAGATGGAGGTCATGCAG 3180
Db 3782 AGTCTGTGGAGGAAGAGATGCAAAATACAGTTTCGAGAGCACAGATGGAGGTCATGCAG 3841
QY 3181 GTGGAATCTTCAACAGATCAATATTTCTCAAGATTCAGAAGTTTGTAAACAAGAACTAT 3240
Db 3842 GTGGAATCTTCAACAGATCAATATTTCTCAAGATTCAGAAGTTTGTAAACAAGAACTAT 3901
QY 3241 GGAAGATACACTCCACGAGAAAAAGAGTTTCTGAAGAAAAACCAACCATGCAAAATG 3300
Db 3902 GGAAGATACACTCCACGAGAAAAAGAGTTTCTGAAGAAAAACCAACCATGCAAAATG 3961
QY 3301 AACGAATGCTATTTTCACTGCTCTCTTTTGTGAATCAATATTCACAAAGGCTTTGATG 3360
Db 3962 AACGAATGCTATTTTCACTGCTCTCTTTTGTGAATCAATATTCACAAAGGCTTTGATG 4021
QY 3361 AAAGCATGCTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGTGAAACTCTT 3420
Db 4022 AAAGCATGCTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGTGAAACTCTT 4081
QY 3421 CCAAAAGCAATCAATATGATATGGAATTTGGAGGAGTACTGGGTCCAGTTCACAAAG 3480
Db 4082 CCAAAAGCAATCAATATGATATGGAATTTGGAGGAGTACTGGGTCCAGTTCACAAAG 4141
QY 3481 ACAGATCTTTTACATTTCCACAGCAGCTGCTCTTTTCCGGGTAACTTTGGCAAGT 3540
Db 4142 ACAGATCTTTTACATTTCCACAGCAGCTGCTCTTTTCCGGGTAACTTTGGCAAGT 4201
QY 3541 CTTTCTGAGTTCAGTGCATGAAATGGCAATTCCTCCTCAGGTCACTCAGTCA 3600
Db 4202 CTTTCTGAGTTCAGTGCATGAAATGGCAATTCCTCCTCAGGTCACTCAGTCA 4261
QY 3601 CTGGTAGGCCCCAGTGTAAATGGCTAGCAATTAGCTCAATATGTTATTTACAGAGGAGAAC 3660
Db 4262 CTGGTAGGCCCCAGTGTAAATGGCTAGCAATTAGCTCAATATGTTATTTACAGAGGAGAAC 4321
QY 3661 AGGCTTATCCTGAGTATTTAACTTACCAGATATGAGGCTCAAGGTATGGTGCATG 3720
Db 4322 AGGCTTATCCTGAGTATTTAACTTACCAGATATGAGGCTCAAGGTATGGTGCATG 4381
QY 3721 GATAAATAGTATTTAAGAACTAATTCCTGAACTGAACCTTAAATCATCAACAGCAGTGT 3780
Db 4382 GATAAATAGTATTTAAGAACTAATTCCTGAACTGAACCTTAAATCATCAACAGCAGTGT 4441
QY 3781 GCCTCTACGTTTACTCCTTTTGTGTAAGAAAAA 3813
Db 4442 GCCTCTACGTTTACTCCTTTTGTGTAAGAAAAA 4474

RESULT 6
AF438201
LOCUS AF438201 6189 bp mRNA linear PRI 12-DEC-2001

DEFINITION Homo sapiens tankyrase II mRNA, complete cds.
ACCESSION AF438201
VERSION AF438201.1 GI:17530294
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
MAMMALIA: Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 6189)
AUTHORS Yin,Y. and Gelmann,E.P.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-2001) Lombardi Cancer Center, Georgetown
University, 3800 Reservoir Rd, NW, Washington, DC 20007, USA
FEATURES
Source
1..6189
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q24"
380..3880
CDS
/note="telomerase binding protein"
/codon_start=1
/product="tankyrase II"
/protein_id="AA140795.1"
/db_xref="GI:17530295"
/translation="MSGRRCAGGGAACAAAEVPAARELFECRNGDVERKRLV
TPEKNSRDYTAGRSTPLHFAAGFGKDVVYLLONGANVQARDGGLTPLHNACSGF
HAEVNILLRHGADPNARDNNYTPLEAAIKGKIDVCIVLLOHGAEPTRITDGRTA
LDLADPSAKAVLTGEYKDELLESARSGNEKMMALLTPLVNCHASDGRKSTPLHLA
AGNRVKIVQLLQHGADVHAKDGLVPLHNACSYGHVETVELLVKHGACVYNDMLW
QPTPLHEAASKNRVEVCSLLSYGADPTLLNCHNSAIDLAPTQLKERLAYEFKHS
LLQAAREADVTRIKKHLSEVMFKHPOTHTALHCAASAPYKPKQICELLKRGAN
INEKTEFLTPLHVAEKANDVVEVVKHAKVNALDNLGOTSLHRAAYCGHLOTCT
LLLSVYCDPNIISLOGFTALOMNENVOQLLEGISLGNSEADRLLEAAKAGDVEY
KKLCTVQVNCRDIEGRQSTPLHFAAGNRVSVVEYLQHGADVHAKDGLVPLHLA
SGHYEVAELLVKHGAVNVADLWKFPTPLHEAAKGYEICKLLQHGADPTKKNRD
GNTPLDLVDKGGTDIOLLRGDAALLDAKGLARVKLLSPDNCNRDTPKRRSTP
LHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDAALLIKYNACVNA
TDKWAFTPLHEAAKQRTQICALLAHGADPTLKNQEGOTPLDLVSADIVSALLTAAM
PPSAPSCVKPQVLNVRSPGATADALSSGSSPSLSAASSLNLDSGFSLSVSVS
SSGTSGSCLEKKEKVPYDFSTIQFVRLNGLHLMIDIPEREQITLDLVEMGHKELKE
TGINAYGHRHKLKIGVRLISQGLNPLYLTUNSGSTILLIDLPDKEFQFQVEEEM
QSTVREHRDGGHAGGFNRYNLTOKVNCNKLWERYTHRRKVEYSEENHNHARMLEF
HGSFVNATIIHKFDEHAYTGMFGAGIYFAENSQSNKYVYGGTGGTGPCVHRDLS
CYICHRQLLFVRTPGLKFLQFSAMKMAHSPPHSHSVTGRPSVNLALAEAYVIYRGEQ
AYPELITYQIMRPEGMDG"
BASE COUNT 1772 a 1233 c 1386 g 1798 t
ORIGIN

Query Match 99.4%; Score 3791.4; DB 9; Length 6189;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 1 CGCGCTGCTCCGCCCGCGGGGCGAGCGGGGGGCGAGGCCCGAGGGCGCGCG 60
Db 157 CGCGCTGCTCCGCCCGCGGGGCGAGCGGGGGGCGAGGCCCGAGGGCGCGCG 216
QY 61 TGGGGCGCGGCCCATGGGACTGCGCGCGGATCCCGGTGACAGCAGGAGCCAGCGCGCGG 120
Db 217 TGGGGCGCGG-CCATGGGACTGCGCGGATCCCGGTGACAGCAGGAGCCAGCGCGCGG 275
QY 121 CCCTGAGCGCGTCTTCTCCGGGGGGCGCTCGCCCTCTCTGTCGCGGGGGCGGGGCTCTCTG 180
Db 276 CCCTGAGCGCGTCTTCTCCGGGGGGCGCTCGCCCTCTCTGTCGCGGGGGCGGGGCTCTCTG 335
QY 181 TCCGGTGTGCTGGCGTGTGCTGGCTGTGGCGCGGCCAGGATCATGTCGGTGCCTGCT 240
Db 336 TCCGGTGTGCTGGCGTGTGCTGGCTGTGGCGCGGCCAGGATCATGTCGGGTCGCGCGCT 395
QY 241 GCGCGCGCGGGGAGCGGCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 396 GCGCGCGCGGGGAGCGGCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 455

Qy 301 CAGAGCTGTTCCAGGCGTGGCGCAACGGGACGTGGAAACAGTCAAGAGGCTGGTGACGC 360
Dd 456 CAGAGCTGTTCCAGGCGTGGCGCAACGGGACGTGGAAACAGTCAAGAGGCTGGTGACGC 515
Qy 361 CTGAGAGGTGAACAGCGCGGACAGCGGCGGCGAGGAATCCACCCCGCTGCACCTTGGCCG 420
Dd 516 CTGAGAGGTGAACAGCGCGGACAGCGGCGGCGAGGAATCCACCCCGCTGCACCTTGGCCG 575
Qy 421 CAGGTTTTGGCGCGAAGACGCTAGTGAATATTTGCTTCAGATGGTGAATGTCACAG 480
Dd 576 CAGGTTTTGGCGCGAAGACGCTAGTGAATATTTGCTTCAGATGGTGAATGTCACAG 635
Qy 481 CACGTGATGATGGGGGCTTATTCCTCTTCATAATGCATGCTCTTTGGTCATGCTGAAG 540
Dd 636 CACGTGATGATGGGGGCTTATTCCTCTTCATAATGCATGCTCTTTGGTCATGCTGAAG 695
Qy 541 TAGTCAATCTCCTTTTGGGACATGGTGAGACGCCCAATGCTCGAGATAATTTGGAATATA 600
Dd 696 TAGTCAATCTCCTTTTGGGACATGGTGAGACGCCCAATGCTCGAGATAATTTGGAATATA 755
Qy 601 CTCCTCTCCATGAAGCTCAATTAAGCAAGATTGATGTTGCTATGCTGCTTACAGC 660
Dd 756 CTCCTCTCCATGAAGCTCAATTAAGCAAGATTGATGTTGCTATGCTGCTTACAGC 815
Qy 661 ATGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATC 720
Dd 816 ATGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATC 875
Qy 721 CATCTGCCAAGCAGTGGCTTACTGGTGAATTAAGAAAGATGAACCTCTTAAGAAAGTGCCTA 780
Dd 876 CATCTGCCAAGCAGTGGCTTACTGGTGAATTAAGAAAGATGAACCTCTTAAGAAAGTGCCTA 935
Qy 781 GGAGTGGCAATCAGAAAAATGATGGCTCTACTCACACCATTAATATCTCAACTGCCACG 840
Dd 936 GGAGTGGCAATCAGAAAAATGATGGCTCTACTCACACCATTAATATCTCAACTGCCACG 995
Qy 841 CAACTGATGGCAAGTCAACTCCATTCATATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Dd 996 CAACTGATGGCAAGTCAACTCCATTCATATTTGGCAGCAGGATATAACAGAGTAAAGA 1055
Qy 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGCCATGCTTAAGATTAAGGTGATCTGG 960
Dd 1056 TTGTACAGCTGTTACTGCAACATGGAGCTGATGCCATGCTTAAGATTAAGGTGATCTGG 1115
Qy 961 TACCATTACAAATGCCGTTCTTATGTCATTTAAGAGTAACCTGAACCTTTTGGTCAAGC 1020
Dd 1116 TACCATTACAAATGCCGTTCTTATGTCATTTAAGAGTAACCTGAACCTTTTGGTCAAGC 1175
Qy 1021 ATGGTGCCTGTGTAATGCAATGGACTTTGGCAATTCACCTCCTTCATGAGCAGCTT 1080
Dd 1176 ATGGTGCCTGTGTAATGCAATGGACTTTGGCAATTCACCTCCTTCATGAGCAGCTT 1235
Qy 1081 CTAAGCAACAGGTTGAAGTATGTTCTCTTCTTAAGTTATGGTGCAGACCCACATGCG 1140
Dd 1236 CTAAGCAACAGGTTGAAGTATGTTCTCTCTCTTAAAGTTATGGTGCAGACCCACATGCG 1295
Qy 1141 TCAATTTGCACAAATAAAGTGCATFAGACTTTGGCTCCCAACACACAGTAAAGAAAGAT 1200
Dd 1296 TCAATTTGCACAAATAAAGTGCATFAGACTTTGGCTCCCAACACACAGTAAAGAAAGAT 1355
Qy 1201 TAGCATATGAATTTAAAGGCCACTGGTTGCTGCAAGCTGCAGGAAGCTGATGTTACTC 1260
Dd 1356 TAGCATATGAATTTAAAGGCCACTGGTTGCTGCAAGCTGCAGGAAGCTGATGTTACTC 1415
Qy 1261 GAATCAAAAAACATCTCTCTGGAATTTGGAATTTCAAGCATCCTCAACACATGAA 1320
Dd 1416 GAATCAAAAAACATCTCTCTGGAATTTGGAATTTCAAGCATCCTCAACACATGAA 1475
Qy 1321 CAGCATTTGATTTGCTGCTGCATCTCCATATCCCAAGAAAGCAAAATATGTTGAACCTGT 1380
Dd 1476 CAGCATTTGATTTGCTGCTGCATCTCCATATCCCAAGAAAGCAAAATATGTTGAACCTGT 1535
Qy 1381 TGCTAAGAAAGGAGCAAAACATCAATGAAAGACTTAAAGAAATTTCTTGACTCCTCTGCACG 1440

Dd 1536 TGCTAAGAAAGGAGCAAAACATCAATGAAAGACTTAAACAATTTCTTGACTCCTCTGCACG 1595
Qy 1441 TGGCATCTCAGAAAAGCTCATATATGATGTTTGAAGTAGTGGTGAACATGAAGCAAGG 1500
Dd 1596 TGGCATCTCAGAAAAGCTCATATATGATGTTTGAAGTAGTGGTGAACATGAAGCAAGG 1655
Qy 1501 TTAATGCTCTGGATAATCTTTGGTCAGACTTCTACACAGAGCTGCATATTTGGTGCATC 1560
Dd 1656 TTAATGCTCTGGATAATCTTTGGTCAGACTTCTACACAGAGCTGCATATTTGGTGCATC 1715
Qy 1561 TACAAACCTGCCGCCCTACTCCTGAGCTATGGGTGATCCTTAACATATATATCCCTTCAGG 1620
Dd 1716 TACAAACCTGCCGCCCTACTCCTGAGCTATGGGTGATCCTTAACATATATATCCCTTCAGG 1775
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGAAATCTACAGCAACTCCTCCAAAGGCTATCT 1680
Dd 1776 GCTTTACTGCTTTACAGATGGGAAATGAAATCTACAGCAACTCCTCCAAAGGCTATCT 1835
Qy 1681 CATTAGGTAAATTCAGAGGACAGACACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCG 1740
Dd 1836 CATTAGGTAAATTCAGAGGACAGACACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCG 1895
Qy 1741 AAACCTGTAAAAAATCTGTACTGTTTCAAGTGTCAACTGCAGAGACATTTGAAGGCGCTC 1800
Dd 1896 AAACCTGTAAAAAATCTGTACTGTTTCAAGTGTCAACTGCAGAGACATTTGAAGGCGCTC 1955
Qy 1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGGAATATCTCG 1860
Dd 1956 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGGAATATCTCG 2015
Qy 1861 TACAGCATGGAGCTGATGTCATGCTTAAAGATAAAGAGGCGCTTGTACCTTTGCACAATG 1920
Dd 2016 TACAGCATGGAGCTGATGTCATGCTTAAAGATAAAGAGGCGCTTGTACCTTTGCACAATG 2075
Qy 1921 CATGTTCTTATGACATTAATGAAGTTGCAGAACTTCTTGTAAACATGAGCAGTAGTTA 1980
Dd 2076 CATGTTCTTATGACATTAATGAAGTTGCAGAACTTCTTGTAAACATGAGCAGTAGTTA 2135
Qy 1981 ATGCTAGCTGATTTATGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATG 2040
Dd 2136 ATGCTAGCTGATTTATGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATG 2195
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAACACAGGATGGAA 2100
Dd 2196 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAACACAGGATGGAA 2255
Qy 2101 ATACTCCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGAG 2160
Dd 2256 ATACTCCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGAG 2315
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAAGGTTGTTTAGCCAGAGTGAAGAACTTGTCTT 2220
Dd 2316 ATGCAGCTTTGCTAGATGCTGCCAAGAAAGGTTGTTTAGCCAGAGTGAAGAACTTGTCTT 2375
Qy 2221 CTCTCATATAATGTAATTTGCCGCGATACCCCAAGCAGACATTTCAACACCTTTACATTTAG 2280
Dd 2376 CTCTCATATAATGTAATTTGCCGCGATACCCCAAGCAGACATTTCAACACCTTTACATTTAG 2435
Qy 2281 CAGCTGGTTATATAATTTTGAAGTTGACAGTATTGTTTACAAACACGAGCTGATGTGA 2340
Dd 2436 CAGCTGGTTATATAATTTTGAAGTTGACAGTATTGTTTACAAACACGAGCTGATGTGA 2495
Qy 2341 ATGCCCAGCAAGAGGAGGACTTATTCCTTTTACATATGTCAGCATCTTACGGCATGTAG 2400
Dd 2496 ATGCCCAGCAAGAGGAGGACTTATTCCTTTTACATATGTCAGCATCTTACGGCATGTAG 2555
Qy 2401 ATGTAGCAGCTCTACTAATAAAGTATTAATGCATGTGTCAATGCCACGACAAATGGGCTT 2460
Dd 2556 ATGTAGCAGCTCTACTAATAAAGTATTAATGCATGTGTCAATGCCACGACAAATGGGCTT 2615
Qy 2461 TCACACCTTTGCGCAAGCAGGCCCAAAAGGAGCAACACAGCTTTTGTGCTTTTGTGCTAG 2520

—

Db	1621	ACAATTGCTGGAAAGCTGCCAAGAGCTGGAGATGTGGAACCTGTAAAAAAAAGCTGTGACTGT	1680
Qy	1766	TCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACACATTCATTTTGCAGC	1825
Db	1681	TCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACACATTCATTTTGCAGC	1740
Qy	1826	TGGGTATAACAGAGTGTCCGTGGTGGAAATATCTGCTACAGATGGAGCTGATGTGCATGC	1885
Db	1741	TGGGTATAACAGAGTGTCCGTGGTGGAAATATCTGCTACAGATGGAGCTGATGTGCATGC	1800
Qy	1886	TAAAGATTAAGAGGAGGCGTTGTACCTTTGCGACAAATGCATCTTCTATATGACATATATGAAGT	1945
Db	1801	TAAAGATTAAGAGGAGGCGTTGTACCTTTGCGACAAATGCATCTTCTATATGACATATATGAAGT	1860
Qy	1946	TGCAGAACTTCTTTGTTTAAACATGGAGCAGTACTGTTAATGTAGCTGATTTATGGAAATTTAC	2005
Db	1861	TGCAGAACTTCTTTGTTTAAACATGGAGCAGTACTGTTAATGTAGCTGATTTATGGAAATTTAC	1920
Qy	2006	ACCTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCA	2065
Db	1921	ACCTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCA	1980
Qy	2066	TGGTGCAGACCCCTACCAAAAAAACAAGGATGGAATATCTCTTTGGATCTTGTAAAGA	2125
Db	1981	TGGTGCAGACCCCTACCAAAAAAACAAGGATGGAATATCTCTTTGGATCTTGTAAAGA	2040
Qy	2126	TGGAGATACAGATATTCAAGATCTGCTTAGGGAGATGCACCTTTGCTAGATGCTGCCAA	2185
Db	2041	TGGAGATACAGATATTCAAGATCTGCTTAGGGAGATGCACCTTTGCTAGATGCTGCCAA	2100
Qy	2186	GAGGGTGTGTTTAGCCAGAGTGAAGAAGTTGCTTCTCCTGATATGTAATATGTCGCGCA	2245
Db	2101	GAGGGTGTGTTTAGCCAGAGTGAAGAAGTTGCTTCTCCTGATATGTAATATGTCGCGCA	2160
Qy	2246	TACCCAAAGCAGACATTCACACCTTTACATTTACGACCTGGTTATATATTTAGAAGT	2305
Db	2161	TACCCAAAGCAGACATTCACACCTTTACATTTAGCAGCTGGTTATATATTTAGAAGT	2220
Qy	2306	TGCAGAGTATTGTTTAAACACGGAGCTGATGTGAATGCCCAAGCAAAAGGAGGACATAT	2365
Db	2221	TGCAGAGTATTGTTTAAACACGGAGCTGATGTGAATGCCCAAGCAAAAGGAGGACATAT	2280
Qy	2366	TCCTTTACATAATGCAGCATTCACGGCATGTAGATGTACAGCTCTACTAATAAAGTA	2425
Db	2281	TCCTTTACATAATGCAGCATTCACGGCATGTAGATGTACAGCTCTACTAATAAAGTA	2340
Qy	2426	TAATGATGTGTCATGCCACGCAAAATGGGCTTTACACCTTTGCACGAAGCAGGCCA	2485
Db	2341	TAATGATGTGTCATGCCACGCAAAATGGGCTTTACACCTTTGCACGAAGCAGGCCA	2400
Qy	2486	AAAGGACGAACACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACTCTTAA	2545
Db	2401	AAAGGACGAACACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACTCTTAA	2460
Qy	2546	AAATCAGGAAGGACAACACCTTTAGATTTAGTTTTCAGCGGATCATGTACAGCGCTCTTCT	2605
Db	2461	AAATCAGGAAGGACAACACCTTTAGATTTAGTTTTCAGCGGATCATGTACAGCGCTCTTCT	2520
Qy	2606	GACAGAGCCATGCCCCCATCTGCTCTGCCCTTTGTTTACAAGCCTCAAGTGTCTCAATGG	2665
Db	2521	GACAGAGCCATGCCCCCATCTGCTCTGCCCTTTGTTTACAAGCCTCAAGTGTCTCAATGG	2580
Qy	2666	TGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTTCAGGTCCATCTAGCCCATCAAG	2725
Db	2581	TGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTTCAGGTCCATCTAGCCCATCAAG	2640
Qy	2726	CCTTTCTGCAGCCAGCAGCTTTGACAACCTTATCTGGGAGTTTTTCAGAACTGCTTTCAGT	2785
Db	2641	CCTTTCTGCAGCCAGCAGCTTTGACAACCTTATCTGGGAGTTTTTCAGAACTGCTTTCAGT	2700
Qy	2786	AGTTAGTTTCAAGTGGACAGAGGGTGTCTTCAAGTTTGGAGAAAAGAGGTTTCCAGAGT	2845
Db	2701	AGTTAGTTTCAAGTGGACAGAGGGTGTCTTCAAGTTTGGAGAAAAGAGGTTTCCAGAGT	2760

QY	2846	AGATTTTAGCATAACTCAATTCGTAAAGGAATCTTGGACCTTTGAGCACCTAATGGATATATT	2905
DB	2761	AGATTTTAGCATAACTCAATTCGTAAAGGAATCTTGGACTTTGAGCACCTAATGGATATATT	2820
QY	2906	TGAGAGAGAACAGATCACATTTGGATGCTATTAGTTGAGATGGGCACAAAGAGCTGAAGGA	2965
DB	2821	TCGAGAGAACAGATCACATTTGGATGCTATTAGTTGAGATGGGCACAAAGAGCTGAAGGA	2880
QY	2966	GATTTGGAATCAATGCTTTATGGACATAGGCACAACTAAATTAAGGAGTGCAGAGACTTAT	3025
DB	2881	GATTTGGAATCAATGCTTTATGGACATAGGCACAACTAAATTAAGGAGTGCAGAGACTTAT	2940
QY	3026	CTCCGGACAAACAGGTCTTTAAACCCATATTAACTTTTCACACCTCTGGTAGTGAACAAT	3085
DB	2941	CTCCGGACAAACAGGTCTTTAAACCCATATTAACTTTTCACACCTCTGGTAGTGAACAAT	3000
QY	3086	TCTTTATAGATCTGTCTCCCTGATGATAAAGAGTTTTCAGTCTGTGGAGGAAGAGATGCAAAAG	3145
DB	3001	TCTTTATAGATCTGTCTCCCTGATGATAAAGAGTTTTCAGTCTGTGGAGGAAGAGATGCAAAAG	3060
QY	3146	TACAGTTCGAGAGCACAGAGATGGAGTCTATGCAGGTGGAATCTTCAACAGATACAAATAT	3205
DB	3061	TACAGTTCGAGAGCACAGAGATGGAGTCTATGCAGGTGGAATCTTCAACAGATACAAATAT	3120
QY	3206	TCTCAAGATTCAGAAGTTTGTAAACAAGAACAATCTATGGGAAAGATACACTCACCGGAGAAA	3265
DB	3121	TCTCAAGATTCAGAAGTTTGTAAACAAGAACAATCTATGGGAAAGATACACTCACCGGAGAAA	3180
QY	3266	ACAAGTCTTCGAAGAAAACCAACCCATGCCAATGAACGAATCTATTTTCATGGGTCTCC	3325
DB	3181	AGAAGTCTTCGAAGAAAACCAACCCATGCCAATGAACGAATCTATTTTCATGGGTCTCC	3240
QY	3326	TTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAGGCATGCGTACATAGTGGTAT	3385
DB	3241	TTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAGGCATGCGTACATAGTGGTAT	3300
QY	3386	GTTTGGAGCTGCAATTTATTTTTCGTGAAAACCTTTCCAAAGCAATCAATATGATATGG	3445
DB	3301	GTTTGGAGCTGCAATTTATTTTTCGTGAAAACCTTTCCAAAGCAATCAATATGATATGG	3360
QY	3446	AAATGGAGGAGTACTGGGTGCTCCAGTTTCACAAAGACAGATCTTGTACATTTGCCACAG	3505
DB	3361	AAATGGAGGAGTACTGGGTGCTCCAGTTTCACAAAGACAGATCTTGTACATTTGCCACAG	3420
QY	3506	GCAGCTGCTCTTTTCCGGGTAACTTGGGAAAGTCTTTCCCTGCAGTTCAGTGCATGAA	3565
DB	3421	GCAGCTGCTCTTTTCCGGGTAACTTGGGAAAGTCTTTCCCTGCAGTTCAGTGCATGAA	3480
QY	3566	AATGCCAATTCCTCCAGGTCACTACCTAGTCACTGGTAGGCCAGTGAATGGCCCT	3625
DB	3481	AATGCCAATTCCTCCAGGTCACTACCTAGTCACTGGTAGGCCAGTGAATGGCCCT	3540
QY	3626	AGCATTAGCTGAATATGTTATTTACAGAGGAGCAACAGGCTTTATCCTGAGTATTAAATTAC	3685
DB	3541	AGCATTAGCTGAATATGTTATTTACAGAGGAGCAACAGGCTTTATCCTGAGTATTAAATTAC	3600
QY	3686	TTACAGATATTAGGCCCTGAAGGTATGATCGATGATAAATAGTTATTTTAAGAACTA	3745
DB	3601	TTACAGATATTAGGCCCTGAAGGTATGATCGATGATAAATAGTTATTTTAAGAACTA	3660
QY	3746	ATTCCACTGAACCTAAAATCATCAAGAGAGAGTGCCTCTACCTGAGTATTCTTGGCTG	3805
DB	3661	ATTCCACTGAACCTAAAATCATCAAGAGAGAGTGCCTCTACCTGAGTATTCTTGGCTG	3720
QY	3806	AAAAAAAAA 3813	
DB	3721	AAAAAAAAA 3728	

DEFINITION Sequence 100 from Patent WO0100849.
ACCESSION AX062241
VERSION AX062241.1 GI:12540142
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4127)
AUTHORS Christenson, E., Demaggio, A.J., Goldman, P.S. and Mcelligott, D.L.
TITLE Tankyrase2 materials and methods
JOURNAL Patent: WO 0100849-A 100 04-JAN-2001;
ICOS CORPORATION (US)
FEATURES
Location/Qualifiers
source 1..4127
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS <2..3511
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC25060.1"
/db_xref="GI:12540143"
/translation="ARMSGRRCAGGGAACASAAAEVPAARELFEARNQDGVK
RLVTPKNSRDTPAGKSTPLPFAAGFGRKDVVEYLLQNGAQAARDGGLPLHNAC
SPGHAENVNLLRHGADPNARNMNTPLHEAAIKGIDKIVLQHGAEPTIRNDG
RTALDLADPSAKAVLGEYKKDELLESARGNEKMMALLPLNVNCHASDGRKSTPL
HLAAGYNRYKIVOLLQHGADVHAKDGLDPLPHNACSVDGHEVTELLYKHGCYNAM
DLWQFPLHEAASKNRVCSLLSYSGADPTLLNCHNSAIDLAFTPQKRLAEYFK
GSHLQARADVTRIKKLSHSEHNVFNKHPQTHETALHCAASPYPKROIICELLRK
GANINKEIKFLPLPHVASEKHNDEYVYVHKEARNALDNLGQTSLHRAAYCGHLQ
TCRLLSYGGDPNIISLQGFALQMGNEVQQLQEGISLGNSEADROLLEAAKAGDV
ETVKKLTGVQNCRDIEGROSTPLPFAAGYNRVSVVEYLLQHGADVHAKDGLDPL
HNACSVDYEAELLVKKHGAENVADLWKEFTPLHEAAKVKYEICKLLQHGADPTK
NRDGNTPDLVQKGDIDQLLGRDAALDAAKGCLARVKLLSPDNVNRCDTQGRH
STPLHUAAGYNLEVAEYLLQHGADVADKGLIPLHNAASYGHVDVAALLIKYNAC
VNATDWAFPLHEAAKGRTOICALLLAHADPTLKNQEGQPLDVLVADVDVSAFLT
AAMPSPALPCYKQVNLGVRSPGATADALSSGSPSSLSAASSLDNISGSFSELLS
VSSSGTSGASSLEKKEVPGVDFSIQFVNLGLEHMDIFEREQITLDVLEMGHKE
LKEGINAYGRHUKLKGVERLISGOGLNPLYLTANTSGGTLIDLSPPDKFEFOSVE
EEMQSTVREHDSGHAGGIFENRYNLIKIQVCNKKLWERYTHRRKEVSENNHANER
MLPFGSPFVNAIHKGFDERHAYIGMFGAGIYFAENSKSNQSYVYGGGGTGPVHK
DRSYCTHRRQLFVRVTLGKSLFQFSAMKMAHSPGPHSHSVTSRPNGLALAEYVYIR
GEQAYPEYLITYQIMRPEGMVDG"
3'UTR 3509..4127
BASE COUNT 1245 a 813 c 934 g 1135 t
ORIGIN
Query Match 94.2%; Score 3595.8; DB 6; Length 4127;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3597; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 215 GGCCAGGATCATGTCGGTCCGCGTGGCGCGGGGGAGCGCGCTGGCGGACGCGCG 274
DB 1 GGCCAGGATCATGTCGGTCCGCGTGGCGCGGGGGAGCGCGCTGGCGGACGCGCG 60
QY 275 GGCCAGGCGTGGAGCGCGCGCGCGAGAGCTGTTCGAGCGCTGCCGCAACGGGACGT 334
DB 61 GGCCAGGCGTGGAGCGCGCGCGCGAGAGCTGTTCGAGCGCTGCCGCAACGGGACGT 120
QY 335 GGAACGAGTCAAGAGCGTGGTGAAGCTGTGAGAGCTGAGAGCTGAGAGCGCGGCGAG 394
DB 121 GGAACGAGTCAAGAGCGTGGTGAAGCTGTGAGAGCTGAGAGCGCGGCGAG 180
QY 395 GAAATCCACCGCGTGCATTCGCGCGAGGTTTGGCGGGAAGACGCTAGTTGAATATT 454
DB 181 GAAATCCACCGCGTGCATTCGCGCGAGGTTTGGCGGGAAGACGCTAGTTGAATATT 240
QY 455 GCTTCAGATGTCGAAATGTCGAGACGCTGATGATGGGCGCTTATTCCTCTTCATAA 514
DB 241 GCTTCAGATGTCGAAATGTCGAGACGCTGATGATGGGCGCTTATTCCTCTTCATAA 300
QY 515 TGCATGCTCTTTTGGTCTGATGAGTCAATCTCTTTTGGGACATGGTGCAGACCC 574
|||||

QY 1655 ACAGCAACTCCTCCAAAGAGGATATCTCATTTAGGTAATTCAGAGGCAGACAGACAATTTGCT 1714
DB 1441 ACAGCAACTCCTCCAAAGAGGATATCTCATTTAGGTAATTCAGAGGCAGACAGACAATTTGCT 1500
QY 1715 GGAAGCTGCAAGAGCTGGAGATGTCGAAACTGTAAAGAACTGTGACTGTTCCAGAGTGT 1774
DB 1501 GGAAGCTGCAAGAGCTGGAGATGTCGAAACTGTAAAGAACTGTGACTGTTCCAGAGTGT 1560
QY 1775 CAACTGCAGAGACATTTGAAGGGCTGAGTCTACACACACTTCATTTTGCAGCTGGGTATTA 1834
DB 1561 CAACTGCAGAGACATTTGAAGGGCTGAGTCTACACACACTTCATTTTGCAGCTGGGTATTA 1620
QY 1835 CAGAGTGTCCGTGGTGGATATATCTGTACAGCATGGAGCTGATGTGCATGCTAAAGATTA 1894
DB 1621 CAGAGTGTCCGTGGTGGATATATCTGTACAGCATGGAGCTGATGTGCATGCTAAAGATTA 1680
QY 1895 AGAGGCCCTGTACCTTTGCACAACTGCATGTTCTTATGACATTAATGAAGTTCGCAACT 1954
DB 1681 AGAGGCCCTGTACCTTTGCACAACTGCATGTTCTTATGACATTAATGAAGTTCGCAACT 1740
QY 1955 TCTTTGTTAAACATGGAGCAGTACTTAATGTAGCTGATTTATGGAATTTTACACCTTTTACA 2014
DB 1741 TCTTTGTTAAACATGGAGCAGTACTTAATGTAGCTGATTTATGGAATTTTACACCTTTTACA 1800
QY 2015 TGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 2074
DB 1801 TGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 1860
QY 2075 CCTTACCAAAAACAGGATGGAATATCTCCTTTTGGATCTTTGTTAAAGATGGAGATAC 2134
DB 1861 CCTTACCAAAAACAGGATGGAATATCTCCTTTTGGATCTTTGTTAAAGATGGAGATAC 1920
QY 2135 AGATATTTCAAGATCTGCTTAGGGAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTG 2194
DB 1921 AGATATTTCAAGATCTGCTTAGGGAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTG 1980
QY 2195 TTTAGCAGAGTGAAGAGTGTCTTCTCCTGATATGTAATTTGCGCGGATACCCAAAG 2254
DB 1981 TTTAGCAGAGTGAAGAGTGTCTTCTCCTGATATGTAATTTGCGCGGATACCCAAAG 2040
QY 2255 CAGACATTTCAACACCTTTACATTTAGCAGCTGTTTATTAATAATTTAGAGTTCACAGTA 2314
DB 2041 CAGACATTTCAACACCTTTACATTTAGCAGCTGTTTATTAATAATTTAGAGTTCACAGTA 2100
QY 2315 TTTGTTTACACACGGAGCTGATGTGAATGCCCAAGCAAAAGGAGGACTTATTCCTTTTACA 2374
DB 2101 TTTGTTTACACACGGAGCTGATGTGAATGCCCAAGCAAAAGGAGGACTTATTCCTTTTACA 2160
QY 2375 TAATGCAGCATCTTAGGGCATGTAGATGTAGCAGCTCTACTAATAAGATTAATGCATG 2434
DB 2161 TAATGCAGCATCTTAGGGCATGTAGATGTAGCAGCTCTACTAATAAGATTAATGCATG 2220
QY 2435 TGTCAATGCCAGGACAAATGGGCTTTACACCTTTGCAGAGCAGCCCAAGAGGAGC 2494
DB 2221 TGTCAATGCCAGGACAAATGGGCTTTACACCTTTGCAGAGCAGCCCAAGAGGAGC 2280
QY 2495 AACACAGCTTTGTGCTTTGCTAGCCCATGGAGTGCAGCCGACTCTTTAAATATCAGCA 2554
DB 2281 AACACAGCTTTGTGCTTTGCTAGCCCATGGAGTGCAGCCGACTCTTTAAATATCAGCA 2340
QY 2555 AGGACAAACACCTTTAGATTTAGTTTACAGCGATGATGTACAGCGCTCTTCTGACAGCAGC 2614
DB 2341 AGGACAAACACCTTTAGATTTAGTTTACAGCGATGATGTACAGCGCTCTTCTGACAGCAGC 2400
QY 2615 CATGCCCATCTGCTCTGCCCTCTTGTACAAAGCCTCAAGTGCATTAATGGTGTGAGAG 2674
DB 2401 CATGCCCATCTGCTCTGCCCTCTTGTACAAAGCCTCAAGTGCATTAATGGTGTGAGAG 2460
QY 2675 CCCAGAGCCACTGCAGATGCTCTCTTTCAGGTCCATCTAGCCCATCAAGCTTTCTGC 2734
DB 2461 CCCAGAGCCACTGCAGATGCTCTCTTTCAGGTCCATCTAGCCCATCAAGCTTTCTGC 2520

QY 2735 AGCCAGCAGTCTTGCACAACTTATCTGGGAGTCTTTCAGAACTGTCTTCAGTAGTTAGTTC 2794
DB 2521 AGCCAGCAGTCTTGCACAACTTATCTGGGAGTCTTTCAGAACTGTCTTCAGTAGTTAGTTC 2580
QY 2795 AAGTGAACAGAGGGTCTTCCAGTTTGGAGAAAAAGGAGTTCCAGGACTAGATTTTAC 2854
DB 2581 AAGTGAACAGAGGGTCTTCCAGTTTGGAGAAAAAGGAGTTCCAGGACTAGATTTTAC 2640
QY 2855 CATTAACCTCAATTCGTAAAGNACTTGGACTTGGACCTTAATGGATATATTTGAGAGAGA 2914
DB 2641 CATTAACCTCAATTCGTAAAGNACTTGGACTTGGACCTTAATGGATATATTTGAGAGAGA 2700
QY 2915 ACAGATCACTTTTGGATGTTATTTAGTTGAGATGGGCAAGGAGCTGAAGGAGATTTGGAAT 2974
DB 2701 ACAGATCACTTTTGGATGTTATTTAGTTGAGATGGGCAAGGAGCTGAAGGAGATTTGGAAT 2760
QY 2975 CAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGCTTATCTCCGAGCA 3034
DB 2761 CAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGCTTATCTCCGAGCA 2820
QY 3035 ACAAGGCTCTTAACCCATATTTAACTTTGAACCTCTGCTAGTGGACAACTTCTTATAGA 3094
DB 2821 ACAAGGCTCTTAACCCATATTTAACTTTGAACCTCTGCTAGTGGACAACTTCTTATAGA 2880
QY 3095 TCTGTCTCTGTATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTGC 3154
DB 2881 TCTGTCTCTGTATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTGC 2940
QY 3155 AGAGCACAGATGGAGTTCATCGAGTGGAACTCTTCAACAGATACAATATTTCTCAAGAT 3214
DB 2941 AGAGCACAGATGGAGTTCATCGAGTGGAACTCTTCAACAGATACAATATTTCTCAAGAT 3000
QY 3215 TCAGAAGTCTTGAACAAAGAACTATGGGAAAGATACACTCACCCGAGAGAAAGAGTTTC 3274
DB 3001 TCAGAAGTCTTGAACAAAGAACTATGGGAAAGATACACTCACCCGAGAGAAAGAGTTTC 3060
QY 3275 TGAAGAAAAACCAACCATGCCCAATCAACGAATGCTATTTTCATGGTCTCCTTTTGTGAA 3334
DB 3061 TGAAGAAAAACCAACCATGCCCAATCAACGAATGCTATTTTCATGGTCTCCTTTTGTGAA 3120
QY 3335 TGCAATATTCACAAAGGCTTTGATGAAAGGCTGCTACATAGGTGGTATGTTTGGAGC 3394
DB 3121 TGCAATATTCACAAAGGCTTTGATGAAAGGCTGCTACATAGGTGGTATGTTTGGAGC 3180
QY 3395 TGGCATTTATTTGCTGAAACTCTTCCAAAGCAATCAATATGATATGGAATTTGGAGG 3454
DB 3181 TGGCATTTATTTGCTGAAACTCTTCCAAAGCAATCAATATGATATGGAATTTGGAGG 3240
QY 3455 AGGTACTGGGTGTCAGTTTCAAAAGACAGATCTTGTTCATTTGCCACAGGAGCTGCT 3514
DB 3241 AGGTACTGGGTGTCAGTTTCAAAAGACAGATCTTGTTCATTTGCCACAGGAGCTGCT 3300
QY 3515 CTTTTCGCGGTAACTTGGGAAAGTCTTCTGCAAGTTCAGTGCATGAAATGGCATTAGC 3574
DB 3301 CTTTTCGCGGTAACTTGGGAAAGTCTTCTGCAAGTTCAGTGCATGAAATGGCATTAGC 3360
QY 3575 TTCTCTCCAGGTCATCTCAGTCTAGTGGCCAGCTGTAATGGCCTAGCATTAGC 3634
DB 3361 TTCTCTCCAGGTCATCTCAGTCTAGTGGCCAGCTGTAATGGCCTAGCATTAGC 3420
QY 3635 TGAATATGTTATTTACAGAGAGACAGGCTTATCTCGATATTTAACTTACCTTACAGAT 3694
DB 3421 TGAATATGTTATTTACAGAGAGACAGGCTTATCTCGATATTTAACTTACCTTACAGAT 3480
QY 3695 TATGAGGCTGAGGTATGTCGATGATAAATAGTATTTTAAAGAACTAACTTCCACATG 3754
DB 3481 TATGAGGCTGAGGTATGTCGATGATAAATAGTATTTTAAAGAACTAACTTCCACATG 3540
QY 3755 AACCTAAAAATCATCAAGCAGCAGTGGCCTCTACGTTTTTACTCCTTTGCTGAAAAAAA 3813
DB 3541 AACCTAAAAATCATCAAGCAGCAGTGGCCTCTACGTTTTTACTCCTTTGCTGAAAAAAA 3599

D	b	1688	GATAAAGRRSCCTTGTAACCTTGCACAAATGCAATGTTCWATATGGACATATTGAAGTTGCA	174
Q	y	1950	GAACTCTCTGTTAAACATGAGCAGTAGTTAATGTAGCTAGCTGATTATATGAAAATTTACACCT	2009
D	b	1748	GAACTCTCTGTTAAACATGAGCAGTAGTTAATGTAGCTAGTTAATGTAGAAAATTTACACCT	1807
Q	y	2010	TTACATGAAGCAGCCAAAGAAGAAATATGAAATTTGCAAACACTTCCTCTCGCATGGT	2069
D	b	1808	TTACATGAAGCAGCCAAAGAAGAAATATGAAATTTGCAAACACTTCCTCTCGCATGGT	1867
Q	y	2070	GCAGACCCTACAAAAAACACAGGATGAAATACTCTCTTTGGATCTTGTTTAAAGATGGA	2129
D	b	1868	GCAGACCCTACAAAAAACACAGGATGAAATACTCTCTTTGGATCTTGTTTAAAGATGGA	1927
Q	y	2130	GATACAGATATTCAAGATCTGTTAGGGAGATGAGAGCTTGTCTAGATGCTGCCAAGAAG	2189
D	b	1928	GATACAGATATTCAAGATCTGTTAGGGAGATGAGAGCTTGTCTAGATGCTGCCAAGAAG	1987
Q	y	2190	GTTTGTGTTAGCCAGATGAAAGATTTGCTCTCTCTGATATGTAATGTAATGCGCGGATACC	2249
D	b	1988	AGTTGTTTAGCCAGATGAAAGATTTGCTCTCTCTGATATGTAATGTAATGCGCGGATACC	2047
Q	y	2250	CAAGGCAGACATTCACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA	2309
D	b	2048	CAAGGCAGACATTCACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA	2107
Q	y	2310	CAGTATTGTTTACAACACGGAGCTGATGTAATGCCCAAGACAAAGAGGACTTATTCCCT	2369
D	b	2108	GAGTATTGTTTACAACACGGAGCTGATGTAATGCCCAAGACAAAGAGGACTTATTCCCT	2167
Q	y	2370	TTACATTAATCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGATATAAT	2429
D	b	2168	TTACATTAATCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGATATAAT	2227
Q	y	2430	GCATGTCTCAATGCCAGGACAAATGGGCTTTTACACCTTTGCCAGAGCAGCCCCAAAG	2489
D	b	2228	GCATGTCTCAATGCCAGGACAAATGGGCTTTTACACCTTTGCCAGAGCAGCCCCAAAG	2287
Q	y	2490	GGAGAACACACGCTTTGTGCTTTGTGTAGCCCATGGAGCTGACCCGACTCTTTAAAAAT	2549
D	b	2288	GGAGAACACACGCTTTGTGCTTTGTGTAGCCCATGGAGCTGACCCGACTCTTTAAAAAT	2347
Q	y	2550	CAGGAAGGACAAACACTTTAGATTTAGTTTTACGGGATGATGCAGGCTTCTCTGACA	2609
D	b	2348	CAGGAAGGACAAACACTTTAGATTTAGTTTTACGGGATGATGCAGGCTTCTCTGACA	2407
Q	y	2610	GCAGCCATGCCCCCATCTGCTCTGCCCTTGTGTACAAGCCTCAAGTCTCAATGGTGTG	2669
D	b	2408	GCAGCCATGCCCCCATCTGCTCTGCCCTTGTGTACAAGCCTCAAGTCTCAATGGTGTG	2467
Q	y	2670	AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTTCAGTCTCATCTAGCCCATCAAGCCTT	2729
D	b	2468	AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTTCAGTCTCATCTAGCCCATCAAGCCTT	2527
Q	y	2730	TCTGCAGCCAGCAGTCTTGCACAACTTATCTGGAGCTTTTTCAGAACATGCTCTTCAGTAGTT	2789
D	b	2528	TCTGCAGCCAGCAGTCTTGCACAACTTATCTGGAGCTTTTTCAGAACATGCTCTTCAGTAGTT	2587
Q	y	2790	AGTTCAAGTGGAAACAGAGGCTGCTTCCAGTTTGGAGAAAAAGAGGTTCCAGGAGTAGAT	2849
D	b	2588	AGTTCAAGTGGAAACAGAGGCTGCTTCCAGTTTGGAGAAAAAGAGGTTCCAGGAGTAGAT	2647
Q	y	2850	TTTAGCATACTCAATTCCTTAAGNACTTTGGACTTTGAGCACCTTAATGGATATATTGAG	2909
D	b	2648	TTTAGCATACTCAATTCCTTAAGNACTTTGGACTTTGAGCACCTTAATGGATATATTGAG	2707
Q	y	2910	AGAAACAGATCACTTTGGATGTATTAGTTGAGATGGGACAAAGAGCTTGAAGGAGATT	2969
D	b	2708	AGAAACAGATCACTTTGGATGTATTAGTTGAGATGGGACAAAGAGCTTGAAGGAGATT	2767
Q	y	2970	GGAACTAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGACTTATCTCC	3029
D	b	2768	GGAACTAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGACTTATCTCC	2827

QY	3030	GGACAACAAGGCTCTTAACCCATATTTAACTTTTGAACACCTCTGGTAGTGGAAACAATTCCTT	3089
Db	2828		2887
		GGACAACAAGGCTCTTAACCCATATTTAACTTTTGAACACCTCTGGTAGTGGAAACAATTCCTT	
QY	3090	ATAGATCTCTCTCCTTGATGATAAAGAGTTTTCAGTCTGTGGAGAAAGAGATGCAAAGTACA	3149
Db	2888		2947
		ATAGATCTCTCTCCTTGATGATAAAGAGTTTTCAGTCTGTGGAGAAAGAGATGCAAAGTACA	
QY	3150	GTTCGAGAGCACAGAGATGGAGTCTATGCAGGTGGAACTCTTCAACAGATACAATATTTCTC	3209
Db	2948		3007
		GTTCGAGAGCACAGAGATGGAGTCTATGCAGGTGGAACTCTTCAACAGATACAATATTTCTC	
QY	3210	AAAGATTGCAAGGTTTGTAAACAGAACTATGCGGAAGATACACTCACCGGAGAAAGAA	3269
Db	3008		3067
		AAAGATTGCAAGGTTTGTAAACAGAACTATGCGGAAGATACACTCACCGGAGAAAGAA	
QY	3270	GTTTCTTGAAGAAAACACACAACCATGCCAATGAACGAATGCTATTTTCATGGGTCTCCTTTT	3329
Db	3068		3127
		GTTTCTGAAACAAACCCACAACCATGCCAATGAACGAATGCTATTTTCATGGGTCTCCTTTT	
QY	3330	GTGAATGCAATTATCCACAAGGCTTTTGATGAAAGCATGCGTACATAGGTGGTATGTTTT	3389
Db	3128		3187
		GTGAATGCAATTATCCACAAGGCTTTTGATGAAAGCATGCGTACATAGGTGGTATGTTTT	
QY	3390	GGAGCTGGCATTTATTTTGGTGAAGACTCTTCCAAAGCAATCAATATGTATATGGAATT	3449
Db	3188		3247
		GGAGCTGGCATTTATTTTGGTGAAGACTCTTCCAAAGCAATCAATATGTATATGGAATT	
QY	3450	GGAGGAGGTACTGGGTGTCAGTTCACAAAGACAGATCTTGTTCACATTTGCCACAGGCAG	3509
Db	3248		3307
		GGAGGAGGTACTGGGTGTCAGTTCACAAAGACAGATCTTGTTCACATTTGCCACAGGCAG	
QY	3510	CTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTCTTCTCGAGTTCAGTCAATGAAAAATG	3569
Db	3308		3367
		CTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTCTTCTCGAGTTCAGTCAATGAAAAATG	
QY	3570	GCACATCTCTCCAGGTGCATCACTCAGTCAAGTCTAGGCCAGTCTAAATGGCCTAGCA	3629
Db	3368		3427
		GCACATCTCTCCAGGTGCATCACTCAGTCAAGTCTAGGCCAGTCTAAATGGCCTAGCA	
QY	3630	TTAGCTGAATATGTTATTTACAGAGAGAACAGGCTTATCCTGAGTATTTAAATTAATCTAC	3689
Db	3428		3487
		TTAGCTGAATATGTTATTTACAGAGAGAACAGGCTTATCCTGAGTATTTAAATTAATCTAC	
QY	3690	CAGATTTATGAGCCCTGAAGGTATGGTCGATGATGAATAAATAGTTATTTTAGAAACTAATTC	3749
Db	3488		3547
		CAGATTTATGAGCCCTGAAGGTATGGTCGATGATGAATAAATAGTTATTTTAGAAACTAATTC	
QY	3750	CAGTGAACCTTAAATCATCAAGCAGCAGTGGCCTTACTGTTTTTACTCCTCTTGGTCAAAA	3809
Db	3548		3607
		CAGTGAACCTTAAATCATCAAGCAGCAGTGGCCTTACTGTTTTTACTCCTCTTGGTCAAAA	
QY	3810	AAAA 3813	
Db	3608		
		AAAA 3611	

RESULT 11
AF264912

LOCUS DEFTNTTTON

DEFINITION
ACCESSION

VERSION
KEYWORDS

SOURCE

ORGANISM

1
2
3
4
5
6
7
8

REFERENCE AUTHORS

ETTER

TITLE

family
 Genes Immun. 2 (1), 52-55 (2001)
 MEDLINE 21190090
 PUBMED 11294570
 REFERENCE 2 (bases 1 to 5810)
 AUTHORS Kuimov, A.N., Kuprash, D.V., Petrov, V.N., Vdovichenko, K.K., Kashkarova, O.A., Boichenko, V.E., Nedospasov, S.A. and Lagarkova, M.A.
 TITLE Direct Submission
 JOURNAL Submitted (08-MAY-2000) Laboratory of Molecular Immunology, A.N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow 119899, Russia
 FEATURES Location/Qualifiers
 source 1..5810
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="10"
 /map="10q22-10q24"
 /tissue_type="breast carcinoma"
 /note="derived from clones MO-BC-203 and 1203"
 <1..5810
 /gene="TNK1"
 /note="a new member of tankyrase gene family; current UniGene cluster Hs.168491"
 1..3501
 /gene="TNK1"
 /note="a new member of tankyrase family"
 /codon_start=1
 /product="tankyrase-like protein"
 /protein_id="AAG44694.1"
 /db_xref="GI:12005976"
 /translation="MSGRCAGGCAACASAAAEVAPAEAREFEACRNGDVERKRVLPKVNRRDAGDPNARDKSTPLHFAAGFGRKDVVYLLQNGANVQARDGGGLIPLHNACSGFHAENVLLRHGADPNARDNWKTPLEHAAIKGKIDVCIVLLOHGAEPTRINDGRTALDADPAKAVLTGEYKDELLESARSNEEKMAALLTPLVNCHASDRKSTPLHLIAAGYNRKIVQLLQHGVDYHAKDGDVPLHNACSGHYEVTELLVHGACVNMADLWQFTPLHAAKNSGVCSLLSYGADPTLLCHINKSAIDIAPTPOLKRLAYEFKSHLLOAREADVTIRKHLSELMVNFKPHOTETALHCAASPYPKRQICELLRLKGANINEXTKEFLPLPHGVTSKANDVYVVKHKAVALNDNLGOTSLHRAAYCGHLQTCRLLSYGDNPILISLQGTALOMGNENYQQLLEGISLGNSEADROLLEAAKAGDVETVKKLCTGVSNCRDLLEGQSTPLHPAAGYNRVSVVEYLLQHGADVHAKDGGGLVPLNACSYGHEVAELLVHGASVNVADLWKFETPLHEAAKGYEICKLQHLQADPTKKNRDGNTPLDVLQDVTDIQDLRLGDALDLDAKKGLRLARVKLSNDPNVNCRDTOGRHSTPLHAAAGYNLEVAHLQHGADVNAQDGGGLIPLHNAASVGHVDVAALLIKYNACVNA TDKWAFTPLHEAAKQGTOLCALILAHGADPTLNKBOGQTPDLVSADVSALLTAAMPSPALPACYPQVNLGVRSPCATADALSSGSPSSLSAASSLDNLGSGSPSELSSVVS SSGTEGASSLKEKVPQVDPSITQFVNRNLGLEHLMDFEREQITLDVLVEMGHKELKE IGINAYGHRKRLKINGBRLISGQQLNPYLTLNTSGSTILILDLSPQDFQSVVEEM QSTVREHRDGHAGGIFNRYNLIKQKCNKWLERYTHRRKEVSEENHNHNERMLF HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSKSNQYVYIGGGTCPPVHKDRS CYICHRQLLCRVTLGRKSLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVYIRGEQ AYPELYITQIMRPEGHVDS"
 complement(1132..1277)
 /note="SHGC-53837"
 complement(3665..3975)
 /note="WI-16054"
 complement(5544..5777)
 /note="WI-8339"
 BASE COUNT 1736 a 1094 c 1231 g 1749 t
 ORIGIN

Query Match 94.0%; Score 3585.8; DB 9; Length 5810;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 225 ATGTCGGTCCGCGTGGCGGGGGAGGCGGCTCGCGAGCGCGCGCGCGAGGCC 284
 Db 1 ATGTCGGTCCGCGTGGCGGGGGAGGCGGCTCGCGAGCGCGCGCGCGAGGCC 60
 Qy 285 GTGGAGCGCGCGCGAGAGCTGTTTCGAGGCGTCCGCCCAACGGGAGCTGGAACGAGTC 344
 Db 61 GTGGAGCGCGCGCGAGAGCTGTTTCGAGGCGTCCGCCCAACGGGAGCTGGAACGAGTC 120

Qy 345 AAGAGCGTGGTACGCGCTGAGAAGGTGAACAGCGCGGACACGCGGCGAGGAATCCACC 404
 Db 121 AAGAGCGTGGTACGCGCTGAGAAGGTGAACAGCGCGGACACGCGGCGAGGAATCCACC 180
 Qy 405 CCCTGCACCTTCGCGCGAGGTGTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAA 464
 Db 181 CCCTGCACCTTCGCGCGAGGTGTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAA 240
 Qy 465 GGTGCAATGTCGAAGCACGTGATGATGGGGCTTATTCTCTCTTCATATGATGCTCT 524
 Db 241 GGTGCAATGTCGAAGCACGTGATGATGGGGCTTATTCTCTCTTCATATGATGCTCT 300
 Qy 525 TTTGGTCATGCTCGAAGTAGTCAATCTCTTTTGGGACATGCTGCAGACCCCAATGCTCGA 584
 Db 301 TTTGGTCATGCTCGAAGTAGTCAATCTCTTTTGGGACATGCTGCAGACCCCAATGCTCGA 360
 Qy 585 GATAATTTGGAATTTATCTCTCTCCATGAAGCTGCAATTAAGGAAGAGATTGATGTTTGC 644
 Db 361 GATAATTTGGAATTTATCTCTCTCCATGAAGCTGCAATTAAGGAAGAGATTGATGTTTGC 420
 Qy 645 ATTGTGCTGTTACAGCATGAGCTGAGCCAAACCATCGGAATACAGATGGAAGGACAGCA 704
 Db 421 ATTGTGCTGTTACAGCATGAGCTGAGCCAAACCATCGGAATACAGATGGAAGGACAGCA 480
 Qy 705 TTGGATTTAGCATCATCTGCCAAGCAGTCTTACTGCTGTAATATAGAAGATGAA 764
 Db 481 TTGGATTTAGCATCATCTGCCAAGCAGTCTTACTGCTGTAATATAGAAGATGAA 540
 Qy 765 CTCTTAGAAGTGGCAGGAGTGGCAATGAAGAAAATGATGCTCTACTCACACCATTA 824
 Db 541 CTCTTAGAAGTGGCAGGAGTGGCAATGAAGAAAATGATGCTCTACTCACACCATTA 600
 Qy 825 AATGTCAACTGCCACGCAAGTGTGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 884
 Db 601 AATGTCAACTGCCACGCAAGTGTGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 660
 Qy 885 TATAACAGAGTAAGATTGACAGCTGTTTACTGCAACATGGAGCTGATGTCCATGCTAA 944
 Db 661 TATAACAGAGTAAGATTGACAGCTGTTTACTGCAACATGGAGCTGATGTCCATGCTAA 720
 Qy 945 GATAAGGTGATGTGTACCATTTACACAATGCCTGTTCTTATGGTCAATTTAAGTAAC 1004
 Db 721 GATAAGGTGATGTGTACCATTTACACAATGCCTGTTCTTATGGTCAATTTAAGTAAC 780
 Qy 1005 GAACCTTTGCTCAAGCATGTGCTGTGTAATGCAATGACCTTTGGGCAATTTCACTCCT 1064
 Db 781 GAACCTTTGCTCAAGCATGTGCTGTGTAATGCAATGACCTTTGGGCAATTTCACTCCT 840
 Qy 1065 CTTTCATGAGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTCTTTAAGTTATGGT 1124
 Db 841 CTTTCATGAGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTCTTTAAGTTATGGT 900
 Qy 1125 GCAGCCCAACACTGCTCAATTTGCACAATAAAAGTGCTATAGACTGGCTCCACACCA 1184
 Db 901 GCAGCCCAACACTGCTCAATTTGCACAATAAAAGTGCTATAGACTGGCTCCACACCA 960
 Qy 1185 CAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCGACCTGTTGCTGCAAGCTGCACGA 1244
 Db 961 CAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCGACCTGTTGCTGCAAGCTGCACGA 1020
 Qy 1245 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTGTAAGTGGTGAATTTCAAGCAT 1304
 Db 1021 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTGTAAGTGGTGAATTTCAAGCAT 1080
 Qy 1305 CCTCAACACATGAACAGCATTTGCATTTGCTGCTGCTCATCTCCATATCCCAAGAAAG 1364
 Db 1081 CCTCAACACATGAACAGCATTTGCATTTGCTGCTGCTCATCTCCATATCCCAAGAAAG 1140
 Qy 1365 CAAATATGTAACCTGTTTGAAGAAAGGACAAACATCAATGAAAGACTTAAGAAATTC 1424
 Db 1141 CAAATATGTAACCTGTTTGAAGAAAGGACAAACATCAATGAAAGACTTAAGAAATTC 1200
 Qy 1425 TTGACTCTCTGACGCTGGCATCTGAGAAAGCTCATAAATGATGTTGTTGAAGTAGTGGTG 1484

QY 1415 TAAAGAAATCTTGACTCCTCTGACAGTGGCATCTGAGAAAGCTCATAAATGATGTTGTGA 1474
Db 1201 TAAAGAAATCTTGACTCCTCTGACAGTGGCATCTGAGAAAGCTCATAAATGATGTTGTGA 1260
QY 1475 AGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTGGATAAATCTTGGTCAGACTTCTCT 1534
Db 1261 AGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTGGATAAATCTTGGTCAGACTTCTCT 1320
QY 1535 ACACAGAGCTGCATATATGCTGCTACAAACCTGCGCGCTACTCCTGAGCTATGGGTG 1594
Db 1321 ACACAGAGCTGCATATATGCTGCTACAAACCTGCGCGCTACTCCTGAGCTATGGGTG 1380
QY 1595 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTTACAGATGGGAAATGAAAATGT 1654
Db 1381 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTTACAGATGGGAAATGAAAATGT 1440
QY 1655 ACAGCAACTCCTCCAGAGAGGTATCTCAATPAGGTAAATTCAGAGCGACAGACAATGTCT 1714
Db 1441 ACAGCAACTCCTCCAGAGAGGTATCTCAATPAGGTAAATTCAGAGCGACAGACAATGTCT 1500
QY 1715 GGAAGCTGCAAGGCTGGAGATGCTGAAACTGTAAAAAACTGTGACTCTTCAGAGTGT 1774
Db 1501 GGAAGCTGCAAGGCTGGAGATGCTGAAACTGTAAAAAACTGTGACTCTTCAGAGTGT 1560
QY 1775 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACCTTCATTTGCAAGCTGGGTATA 1834
Db 1561 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACCTTCATTTGCAAGCTGGGTATA 1620
QY 1835 CAGAGTCCGCTGGTGAATATCTGCTACAGCATGAGCTGATGTCATGCTAAAGATA 1894
Db 1621 CAGAGTCCGCTGGTGAATATCTGCTACAGCATGAGCTGATGTCATGCTAAAGATA 1680
QY 1895 AGAGGCTTGTACCTTTGCAATGCATGCTTCTTATGACATTAATGAAGTTCGAGAACT 1954
Db 1681 AGAGGCTTGTACCTTTGCAATGCATGCTTCTTATGACATTAATGAAGTTCGAGAACT 1740
QY 1955 TCTTGTAAACATGAGCAGTAGTAAATGATGCTGATTTATGAGAAATTTACACCTTTTACA 2014
Db 1741 TCTTGTAAACATGAGCAGTAGTAAATGATGCTGATTTATGAGAAATTTACACCTTTTACA 1800
QY 2015 TGAAGCAGCAGCAAAAAGGAAATATGAAATTTGCAAACTTCTCTCCAGCATGGTCGAGA 2074
Db 1801 TGAAGCAGCAGCAAAAAGGAAATATGAAATTTGCAAACTTCTCTCCAGCATGGTCGAGA 1860
QY 2075 CCTACCAAAAAACAGGATGGAATACTCCTTTGGATCTTTGTTAAAGATGGAGATAC 2134
Db 1861 CCTACCAAAAAACAGGATGGAATACTCCTTTGGATCTTTGTTAAAGATGGAGATAC 1920
QY 2135 AGATATTCAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGGTTG 2194
Db 1921 AGATATTCAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGGTTG 1980
QY 2195 TTTAGCCAGAGTGAAGAAGTTGCTTCTCTGATAATGTAATTTGCCGCGATACCCAAGG 2254
Db 1981 TTTAGCCAGAGTGAAGAAGTTGCTTCTCTGATAATGTAATTTGCCGCGATACCCAAGG 2040
QY 2255 CAGACATTCACACCTTTACATTTAGCAGCTGGTTAATAATTTAGAGTTGCCAGATGA 2314
Db 2041 CAGACATTCACACCTTTACATTTAGCAGCTGGTTAATAATTTAGAGTTGCCAGATGA 2100
QY 2315 TTTTGTACAAACAGGAGCTGATGTAATGCCAAGCAAGAGGACTTTATTCCTTTTACA 2374
Db 2101 TTTTGTACAAACAGGAGCTGATGTAATGCCAAGCAAGAGGACTTTATTCCTTTTACA 2160
QY 2375 TAATGCAGCATCTTACGGGCGATGTAGATGAGCAGCTCTACTAATAAAGATAATATGTCATG 2434
Db 2161 TAATGCAGCATCTTACGGGCGATGTAGATGAGCAGCTCTACTAATAAAGATAATATGTCATG 2220
QY 2435 TGTCAATGCCCGGCAAAATGGGCTTTACACCTTTGACAGAGCAGGCCAAAGGGAGG 2494
Db 2221 TGTCAATGCCCGGCAAAATGGGCTTTACACCTTTGACAGAGCAGGCCAAAGGGAGG 2280

QY 2495 AACACAGCTTTGTGCTTTGTTGCTAGCCCTGAGAGCTGACCCGAGCTCTTAAAAATCAGGA 2554
Db 2281 AACACAGCTTTGTGCTTTGTTGCTAGCCCTGAGAGCTGACCCGAGCTCTTAAAAATCAGGA 2340
QY 2555 AGGACAAACACCTTTAGATTTAGTTTTCAGGCGATGATGTCAGCGCTCTTCTGACAGCAGC 2614
Db 2341 AGGACAAACACCTTTAGATTTAGTTTTCAGGCGATGATGTCAGCGCTCTTCTGACAGCAGC 2400
QY 2615 CATGCCCCCAATCTGCTCTGCCCTCTTCTTACAAGCCTCAAGTGTCTCAATGGTGTGAGAAG 2674
Db 2401 CATGCCCCCAATCTGCTCTGCCCTCTTCTTACAAGCCTCAAGTGTCTCAATGGTGTGAGAAG 2460
QY 2675 CCAGAGGACCACTGCAGATGCTCTCTTCAGGTGCCATCTAGCCCATCAAGCCCTTTCTGC 2734
Db 2461 CCAGAGGACCACTGCAGATGCTCTCTTCAGGTGCCATCTAGCCCATCAAGCCCTTTCTGC 2520
QY 2735 AGCCAGCAGCTTTGACAACTTATCTGGGAGTCTTTCAGAACTGTCTTCACTAGTATTAGTTC 2794
Db 2521 AGCCAGCAGCTTTGACAACTTATCTGGGAGTCTTTCAGAACTGTCTTCACTAGTATTAGTTC 2580
QY 2795 AAGTGGAAACAGAGGGTCTTCCAGTTTGGAGAAAAGGAGGTTCAGGAGTAGATTTAG 2854
Db 2581 AAGTGGAAACAGAGGGTCTTCCAGTTTGGAGAAAAGGAGGTTCAGGAGTAGATTTAG 2640
QY 2855 CATAACTCAATTCCTAAGGAATCTTGGACTTGGACACTTAATGGATATATTTGAGAGAGA 2914
Db 2641 CATAACTCAATTCCTAAGGAATCTTGGACTTGGACACTTAATGGATATATTTGAGAGAGA 2700
QY 2915 ACAGATCACTTTGGATGATATGATTTGAGATGGGCAACAAGGAGCTGAAGGAGATTGGAAT 2974
Db 2701 ACAGATCACTTTGGATGATATGATTTGAGATGGGCAACAAGGAGCTGAAGGAGATTGGAAT 2760
QY 2975 CAATGCTTATGGACATAGGCACAACTTAATTAAGGAGTGCAGAGACTTATCTCCGAGCA 3034
Db 2761 CAATGCTTATGGACATAGGCACAACTTAATTAAGGAGTGCAGAGACTTATCTCCGAGCA 2820
QY 3035 ACAAGTCTTAAACCCATATTTAACTTTGACACCTCTGCTAGTGGAACTTCTTTATAGA 3094
Db 2821 ACAAGTCTTAAACCCATATTTAACTTTGACACCTCTGCTAGTGGAACTTCTTTATAGA 2880
QY 3095 TCTGCTCTCTGATGATTAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAGTACAGTTGCG 3154
Db 2881 TCTGCTCTCTGATGATTAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAGTACAGTTGCG 2940
QY 3155 AGAGCACAGAGATGGAGTTCATGCAAGTGGAACTTCTTCAACAGATACAATATTTCTCAAGAT 3214
Db 2941 AGAGCACAGAGATGGAGTTCATGCAAGTGGAACTTCTTCAACAGATACAATATTTCTCAAGAT 3000
QY 3215 TCAGAGGTTTGTACAGAACTATGGGAAAGATACACTCACCGGAGAAAAGAGTTTC 3274
Db 3001 TCAGAGGTTTGTACAGAACTATGGGAAAGATACACTCACCGGAGAAAAGAGTTTC 3060
QY 3275 TGAAGAAAACCAACCATGCCAATGAACGATGCTTATTCATGGCTCTCCTTTTGTGAA 3334
Db 3061 TGAAGAAAACCAACCATGCCAATGAACGATGCTTATTCATGGCTCTCCTTTTGTGAA 3120
QY 3335 TGCAATTTATCCACAAGGCTTTGATGAAAAGCATGCTACATAGTGGTATGTTTGGAGC 3394
Db 3121 TGCAATTTATCCACAAGGCTTTGATGAAAAGCATGCTACATAGTGGTATGTTTGGAGC 3180
QY 3395 TGGCAATTTATTTGCTGAAAACCTTCTTCCAAAAGCAATCAATATGTAATGGAATGCGAG 3454
Db 3181 TGGCAATTTATTTGCTGAAAACCTTCTTCCAAAAGCAATCAATATGTAATGGAATGCGAG 3240
QY 3455 AGGTACTGGGTCTGCAGTTTCAAAAGACAGATCTTGTACATTTGCGCAGGCGAGCTGCT 3514
Db 3241 AGGTACTGGGTCTGCAGTTTCAAAAGACAGATCTTGTACATTTGCGCAGGCGAGCTGCT 3300
QY 3515 CTTTTCGCGGTAACTTTGGGAAAGTCTTCTTCTGCAAGTTCAGTGAATGAAATGGCACA 3574
Db 3301 CTTTTCGCGGTAACTTTGGGAAAGTCTTCTTCTGCAAGTTCAGTGAATGAAATGGCACA 3360
QY 3575 TTTCTCTCCAGGTCATCACTCACTGCTAGGTCAGTGGTGGCCCACTGTAATGGCCTAGCATTAGC 3634

QY	1245	GAAGCTGATGTTACTCGAATCAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT	1304
DB	1021		1080
		GAAGCTGATGTTACTCGAATCAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT	
QY	1305	CCTCAACACATGAACAGCATTTGCATTTGTGTGTGTCATCTCCATATCCCAAAAGAAG	1364
DB	1081		1140
		CCTCAACACATGAACAGCATTTGCATTTGTGTGTGTCATCTCCATATCCCAAAAGAAG	
QY	1365	CAATATGTGAAGCTTTGCTTAAGAAAGAGGACAAACATCAATGAAAGACTTAAAGAAATTC	1424
DB	1141		1200
		CAATATGTGAAGCTTTGCTTAAGAAAGAGGACAAACATCAATGAAAGACTTAAAGAAATTC	
QY	1425	TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATATATGATGTTGTTGAAGTACTGGTG	1484
DB	1201		1260
		TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTACTGGTG	
QY	1485	AAACATGAAGCAAAAGGTTAATGCTCTGGATAATCTTTGGTTCAGACTTCTCTACACAGAGCT	1544
DB	1261		1320
		AAACATGAAGCAAAAGGTTAATGCTCTGGATAATCTTTGGTTCAGACTTCTCTACACAGAGCT	
QY	1545	GCATATTTGGTCAATCTACAAACCTGCCGCTACTCTCTGAGCTATFGGCTGTGATCTCTAAC	1604
DB	1321		1380
		GCATATTTGGTCAATCTACAAACCTGCCGCTACTCTCTGAGCTATFGGCTGTGATCTCTAAC	
QY	1605	ATTATATCCCTTCAGGCGTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTC	1664
DB	1381		1440
		ATTATATCCCTTCAGGCGTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTC	
QY	1665	CTCAAGAGGGTATCTCAATAGGTAAATCAGAGGCGAGACAGACAATTTGCTGGAAGCTGCA	1724
DB	1441		1500
		CTCAAGAGGGTATCTCAATAGGTAAATCAGAGGCGAGACAGACAATTTGCTGGAAGCTGCA	
QY	1725	AAGGCTGGAGATGCGAAACTGTFAAAAAAAGCTGTACTGTTTCAGAGTGCAACTGCAGA	1784
DB	1501		1560
		AAGGCTGGAGATGCGAAACTGTFAAAAAAAGCTGTACTGTTTCAGAGTGCAACTGCAGA	
QY	1785	GACATTTGAAGGGCTGAGTCTACACACTTTCATTTTGAGCTGGGTATACAGAGTGTCC	1844
DB	1561		1620
		GACATTTGAAGGGCTGAGTCTACACACTTTCATTTTGAGCTGGGTATACAGAGTGTCC	
QY	1845	GTGTGTGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGCGTT	1904
DB	1621		1680
		GTGTGTGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGCGTT	
QY	1905	GATCCTTTGCACAATCGATGTTCTTATGGACAATTATGAAGTTGAGAACCTTCTTTGTTAAA	1964
DB	1681		1740
		GATCCTTTGCACAATCGATGTTCTTATGGACAATTATGAAGTTGAGAACCTTCTTTGTTAAA	
QY	1965	CATGGACAGTACTTAAATGACTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA	2024
DB	1741		1800
		CATGGACAGTACTTAAATGACTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA	
QY	2025	GCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGTGTCAGACCCCTACCCAA	2084
DB	1801		1860
		GCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGTGTCAGACCCCTACCCAA	
QY	2085	AAAACAGGATGGAATATCTCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTCAA	2144
DB	1861		1920
		AAAACAGGATGGAATATCTCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTCAA	
QY	2145	GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGCTTCTTTAGCCAGA	2204
DB	1921		1980
		GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGCTTCTTTAGCCAGA	
QY	2205	GTGAAGAAGTTGCTTCTCCTTGATTAATGTAATTTGCCGCGATACCCCAAGCAGACATTCA	2264
DB	1981		2040
		GTGAAGAAGTTGCTTCTCCTTGATTAATGTAATTTGCCGCGATACCCCAAGCAGACATTCA	
QY	2265	ACACCTTTACATTTAGCAGCTGGTTATATAATTTAGAAAGTTGAGAGTATTGTTACAA	2324
DB	2041		2100
		ACACCTTTACATTTAGCAGCTGGTTATATAATTTAGAAAGTTGAGAGTATTGTTACAA	
QY	2325	CACGGAGCTGATGTGAATGCCCAAGACAAAGAGGAGCTTATTCTCTTTACATAATGCGACA	2384

[illegible]

Db 3001 TGTACAGAAACATATGGAAAGATACACTACCGGAGAAAAGAGTTTCTGAGAAAC 3060
QY 3285 CACAACCATGCAATGAACGAATGCTATTTTCATGGGTCTCTTTTGTGAATGCAATATC 3344
Db 3061 CACAACCATGCAATGAACGAATGCTATTTTCATGGGTCTCTTTTGTGAATGCAATATC 3120
QY 3345 CACAAGGCTTTGATGAAGGCATCGCTACATAGTGGTATGTTTGGAGCTGGCATTTAT 3404
Db 3121 CACAAGGCTTTGATGAAGGCATCGCTACATAGTGGTATGTTTGGAGCTGGCATTTAT 3180
QY 3405 TTTGCTGAAACTCTTCCAAAGCAATCAATATATATATGAAATGGAGGAGTACTGGG 3464
Db 3181 TTTGCTGAAACTCTTCCAAAGCAATCAATATATATGAAATGGAGGAGTACTGGG 3240
QY 3465 TGTCAGTTTCAAAAGACAGATCTTTGTTACATTTGCCACAGGCAGTCTCTTTTGGCGG 3524
Db 3241 TGTCAGTTTCAAAAGACAGATCTTTGTTACATTTGCCACAGGCAGTCTCTTTTGGCGG 3300
QY 3525 GTAACTTTGGGAAAGCTTTCTTCCAGTTTCACTGAGTTCACTGCAATGAAATGGCACATCTCTCCCA 3584
Db 3301 GTAACTTTGGGAAAGCTTTCTTCCAGTTTCACTGAGTTCACTGCAATGAAATGGCACATCTCTCCCA 3360
QY 3585 GGTATCACTCACTGACCTGAGTGGAGCCAGTCTGTAATGGCTTAGCATTAGCTGAATATGTT 3644
Db 3361 GGTATCACTCACTGACCTGAGTGGAGCCAGTCTGTAATGGCTTAGCATTAGCTGAATATGTT 3420
QY 3645 ATTTACAGAGGAGACAGGCTTATCTCTGATTTAATTAATTTACCTACCATATGAGGCT 3704
Db 3421 ATTTACAGAGGAGACAGGCTTATCTCTGATTTAATTAATTTACCTACCATATGAGGCT 3480
QY 3705 GAAGTATGGTCTGATGGA 3722
Db 3481 GAAGTATGGTCTGATGGA 3498

RESULT 15
LOCUS AX133694 3394 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0130987.
ACCESSION AX133694
VERSION AX133694.1 GI:14139712
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3394)
AUTHORS Luo, Y., Chan, E., Xu, X. and Huang, B.
TITLE Tankyrase h, compositions involved in the cell cycle and methods of use
JOURNAL Patent: WO 0130987-A 1 03-MAY-2001;
Rigel Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source 1..3394
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1046 a 662 c 767 g 919 t
ORIGIN
Query Match 88.8%; Score 3387.6; DB 6; Length 3394;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3390; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 423 GGTGTTGGCGGAAAGACGATGTTGAATATTTGCTTCAGAAATGGTGCATAATGCAAGCA 482
Db 1 GGTGTTGGCGGAAAGACGATGTTGAATATTTGCTTCAGAAATGGTGCATAATGCAAGCA 60
QY 483 CGTGATGATGGGGCCCTTATCTCTTCATATGATGCTCTTTTGGTCATGCTGAAGTA 542
Db 61 CGTGATGATGGGGCCCTTATCTCTTCATATGATGCTCTTTTGGTCATGCTGAAGTA 120
QY 543 GTCAATCTCTTTTGGGACATGGTGCAGACGCCCAATGCTCGAGATAATGGAATTATACT 602

Db 121 GTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATACT 180
QY 603 CCTCTCCATGAAGCTGCAATTTAAAGAAAGATTTGATTTGCTGCTGTTTACAGCAT 562
Db 181 CCTCTCCATGAAGCTGCAATTTAAAGAAAGATTTGATTTGCTGCTGTTTACAGCAT 240
QY 663 GGAGCTGAGCGCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCCA 722
Db 241 GGAGCTGAGCGCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCCA 300
QY 723 TCTGCCAAAGCAGTGGCTTACTGGTGAATATAGAAAGATGAACTCTTAGAAAGTGCAGG 782
Db 301 TCTGCCAAAGCAGTGGCTTACTGGTGAATATAGAAAGATGAACTCTTAGAAAGTGCAGG 360
QY 783 AGTGGCAATGAAGAAATATGATGGCTCTACTCAGCACCATTAAATGTCACACTGCCACGCA 842
Db 361 AGTGGCAATGAAGAAATATGATGGCTCTACTCAGCACCATTAAATGTCACACTGCCACGCA 420
QY 843 AGTATGGGCAAGGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 902
Db 421 AGTATGGGCAAGGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 480
QY 903 GTACAGCTGTTACTGCAACATGAGCTGATGTCATCTTAAAGATAAAGGTGATCTGGTA 962
Db 481 GTACAGCTGTTACTGCAACATGAGCTGATGTCATCTTAAAGATAAAGGTGATCTGGTA 540
QY 963 CCATTACACAATGGCTGTTCTTATGGTCATTTGAAGTAAGTGAAGTCTTGGTCAAGCAT 1022
Db 541 CCATTACACAATGGCTGTTCTTATGGTCATTTGAAGTAAGTGAAGTCTTGGTCAAGCAT 600
QY 1023 GGTGCTGTGTAATGCAATGGACTTGTGCAATTTCACTCTCTCATGAGCAGCTTCT 1082
Db 601 GGTGCTGTGTAATGCAATGGACTTGTGCAATTTCACTCTCTCATGAGCAGCTTCT 660
QY 1083 AAGAACAGGTTGAAGTATGTTCTTCTTCTTAAAGTATGTTGTCAGACCCCAACTGCTC 720
Db 661 AAGAACAGGTTGAAGTATGTTCTTCTTCTTAAAGTATGTTGTCAGACCCCAACTGCTC 780
QY 1143 AATTGTCACATRAAAGTGTATAGACTTGGCTCCACACACAGATTTAAAGAAAGATTA 1202
Db 721 AATTGTCACATRAAAGTGTATAGACTTGGCTCCACACACAGATTTAAAGAAAGATTA 780
QY 1203 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGGAGAGCTGATGTTACTCGA 1262
Db 781 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGGAGAGCTGATGTTACTCGA 840
QY 1263 ATCAAAACACATCTCTCTGGAATGTTGAATTTCAAGCATCTCAACACATGAAGCA 1322
Db 841 ATCAAAACACATCTCTCTGGAATGTTGAATTTCAAGCATCTCAACACATGAAGCA 900
QY 1323 GCATGCAATTTGCTGCTGCAATCTCCATATCCCAAAAGAAAGCAATATGTGAAGTCTTG 1382
Db 901 GCATGCAATTTGCTGCTGCAATCTCCATATCCCAAAAGAAAGCAATATGTGAAGTCTTG 960
QY 1383 CTAAGAAAGAGGCAACATCAATGAAAGACTAAAGAAATTTGACTCTCTGCAAGTG 1442
Db 961 CTAAGAAAGAGGCAACATCAATGAAAGACTAAAGAAATTTGACTCTCTGCAAGTG 1020
QY 1443 GCATCTGAGAAGCTCATAATGATGTTGTTGAAGTGTGGTGAACATGAAGCAAGGTT 1502
Db 1021 GCATCTGAGAAGCTCATAATGATGTTGTTGAAGTGTGGTGAACATGAAGCAAGGTT 1080
QY 1503 AATGCTCTGGATAATCTTGGTCAGACTTCTACACAGAGCTGCATATTTGGTGCATCTA 1562
Db 1081 AATGCTCTGGATAATCTTGGTCAGACTTCTACACAGAGCTGCATATTTGGTGCATCTA 1140
QY 1563 CAAACCTGCCCTTACTCTGAGCTATGGTGTGATCTTAACATTTATATCCCTTCAGGCG 1622
Db 1141 CAAACCTGCCCTTACTCTGAGCTATGGTGTGATCTTAACATTTATATCCCTTCAGGCG 1200
QY 1623 TTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCCAAGAGGTTATCTCA 1682
Db 1623 TTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCCAAGAGGTTATCTCA 1682

Db 1201 TTTACTGCTTACAGATGGGAAATGAAATGTACAGCAACTCTCCAAAGAGGGTATCTCA 1260
Qy 1683 TTAGGTAATTCAGAGGCAGACAGCAATTCCTGGAGCTGCAAGGCTGGAGATGTCGAA 1742
Db 1261 TTAGGTAATTCAGAGGCAGACAGCAATTCCTGGAGCTGCAAGGCTGGAGATGTCGAA 1320
Qy 1743 ACTGTAAAAAAGCTGTGTACTGTTTCAGAGTGTCAACTCGACAGACATGGAAGGGCTGAG 1802
Db 1321 ACTGTAAAAAAGCTGTGTACTGTTTCAGAGTGTCAACTCGACAGACATGGAAGGGCTGAG 1380
Qy 1803 TCTACACCATTCTATTTGACAGCTGGGTATTAACAGAGTGTCCGTGGTGGATATCTGCTA 1862
Db 1381 TCTACACCATTCTATTTGACAGCTGGGTATTAACAGAGTGTCCGTGGTGGATATCTGCTA 1440
Qy 1863 CAGCATGAGCTGATGTGCATGCTAAAGATAAAGAGGSCCTGTACCTTTTCACAATGCA 1922
Db 1441 CAGCATGAGCTGATGTGCATGCTTAAGATAAAGAGGSCCTGTACCTTTTCACAATGCA 1500
Qy 1923 TGTTCCTATGAGCATATTAAGAGTTGCAAGACTTCTTGTAAACATGGAGCAGTAGTTAAT 1982
Db 1501 TGTTCCTATGAGCATATTAAGAGTTGCAAGACTTCTTGTAAACATGGAGCAGTAGTTAAT 1560
Qy 1983 GTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 2042
Db 1561 GTAGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 1620
Qy 2043 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCRAAAAAACAGGGATGGAAT 2102
Db 1621 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCRAAAAAACAGGGATGGAAT 1680
Qy 2103 ACTCCTTTGATCTTGTAAAGATGGAGATACAGATATTCAGAGATCTGCTTAGGGAGAT 2162
Db 1681 ACTCCTTTGATCTTGTAAAGATGGAGATACAGATATTCAGATATCTGCTTAGGGAGAT 1740
Qy 2163 GCAGCTTTGATCTGCTGCCAGAGGGTGTATTAAGAGTGTGTTTACCAGAGTGAAGATGCTCTCT 2222
Db 1741 GCAGCTTTGATCTGCTGCCAGAGGGTGTATTAAGAGTGTGTTTACCAGAGTGAAGATGCTCTCT 1800
Qy 2223 CCTGATAATGTAATTCGCCGATACCAAGGCAGACATTCACACCTTTACATTTAGCA 2282
Db 1801 CCTGATAATGTAATTCGCCGATACCAAGGCAGACATTCACACCTTTACATTTAGCA 1860
Qy 2283 GCTGGTTATTAATTTAGAGTTGACAGATTTGTTTACACACGAGCTGTATGTTGAAT 2342
Db 1861 GCTGGTTATTAATTTAGAGTTGACAGATTTGTTTACACACGAGCTGTATGTTGAAT 1920
Qy 2343 GCCAAGACAAAGGAGACTTATTCCTTTTACATAATGACAGCTTACGGGATGTAGAT 2402
Db 1921 GCCAAGACAAAGGAGACTTATTCCTTTTACATAATGACAGCTTACGGGATGTAGAT 1980
Qy 2403 GTAGCAGCTCTACTAATAAGTATAATGTCATGTCTCAATGCCAGCAAAATGGGCTTTC 2462
Db 1981 GTAGCAGCTCTACTAATAAGTATAATGTCATGTCTCAATGCCAGCAAAATGGGCTTTC 2040
Qy 2463 ACACCTTTGACCAAGCAGCCAAAAGGGAGCAGACAGCTTTGCTGTTGCTAGCC 2522
Db 2041 ACACCTTTGACCAAGCAGCCAAAAGGGAGCAGACAGCTTTGCTGTTGCTAGCC 2100
Qy 2523 CATGGAGCTGACCGACTCTTAAAAATCAGGAAGGCAACACCTTTAGATTTAGTTTCA 2582
Db 2101 CATGGAGCTGACCGACTCTTAAAAATCAGGAAGGCAACACCTTTAGATTTAGTTTCA 2160
Qy 2583 GCGGATGATGTACAGCTCTCTTGACAGCAGCCATGCCCCACTCTCTGCTGCTCTTGT 2642
Db 2161 GCGGATGATGTACAGCTCTCTTGACAGCAGCCATGCCCCACTCTCTGCTGCTCTTGT 2220
Qy 2643 TACAAGCTCAAGTGTCAATGGTGTGAGAGCCAGAGCCACTGCAGATGCTCTCTCT 2702
Db 2221 TACAAGCTCAAGTGTCAATGGTGTGAGAGCCAGAGCCACTGCAGATGCTCTCTCT 2280
Qy 2703 TCAGGTCCTATAGCCCATCAAGCCCTTTCTGACAGCAGCAGCTTTCGACAACTTATCTGGG 2762
Db 2281 TCAGGTCCTATAGCCCATCAAGCCCTTTCTGACAGCAGCAGCTTTCGACAACTTATCTGGG 2340

Qy 2763 AGTTTTTCAGAACTGCTCTCAGTAGTCTAGTCAAGTGAACAGAGGGTCTTCCAGTTTG 2822
Db 2341 AGTTTTTCAGAACTGCTCTCAGTAGTCTAGTCAAGTGAACAGAGGGTCTTCCAGTTTG 2400
Qy 2823 GAGAAAAAGGAGGTTCCAGGAGTAGATTTAGCAATAACTCAATTCGTAAAGAACTTTGGA 2882
Db 2401 GAGAAAAAGGAGGTTCCAGGAGTAGATTTAGCAATAACTCAATTCGTAAAGAACTTTGGA 2460
Qy 2883 CTTGAGCACCTTAATGGATATATTTGAGAGAGAACAGATCACTTTTGATGTATTTAGTTGAG 2942
Db 2461 CTTGAGCACCTTAATGGATATATTTGAGAGAGAACAGATCACTTTTGATGTATTTAGTTGAG 2520
Qy 2943 ATGGGGCACAGGAGCTGAAGGAGATTTGGAATCAANTGCTTATGGACATAGGACAAACTA 3002
Db 2521 ATGGGGCACAGGAGCTGAAGGAGATTTGGAATCAANTGCTTATGGACATAGGACAAACTA 2580
Qy 3003 ATTAAGAGGTCGAGAGACTTATCTCCGACAAAGAGTCTTAACCCATATTTAACTTTG 3062
Db 2581 ATTAAGAGGTCGAGAGACTTATCTCCGACAAAGAGTCTTAACCCATATTTAACTTTG 2640
Qy 3063 AACACCTCTGCTAGTGAACAAATTTCTATAGATCTGCTCTCTGATGATTAAGAGTTTCAG 3122
Db 2641 AACACCTCTGCTAGTGAACAAATTTCTATAGATCTGCTCTCTGATGATTAAGAGTTTCAG 2700
Qy 3123 TCTGTGAGGAGAGATGCAAGATACAGTTCGAGAGCACAGATGGAGGTCATGCAGGT 3182
Db 2701 TCTGTGAGGAGAGATGCAAGATACAGTTCGAGAGCACAGATGGAGGTCATGCAGGT 2760
Qy 3183 GGAATCTTCAACAGATACAATATTTCTCAAGATTCAGAAAGTTCGTAACAAAGAACTATGG 3242
Db 2761 GGAATCTTCAACAGATACAATATTTCTCAAGATTCAGAAAGTTCGTAACAAAGAACTATGG 2820
Qy 3243 GAAAGATACACTCACCGGAGAAAGAGTTTCTGAAGAAAAACCAACCATGCGCAATGAA 3302
Db 2821 GAAAGATACACTCACCGGAGAAAGAGTTTCTGAAGAAAAACCAACCATGCGCAATGAA 2880
Qy 3303 CGAATGCTATTTCAATGCTCTCTCTTTGTAATGCAATTTATCCACAAAGGCTTTGATGAA 3362
Db 2881 CGAATGCTATTTCAATGCTCTCTCTTTGTAATGCAATTTATCCACAAAGGCTTTGATGAA 2940
Qy 3363 AGGCATCGTACATAGGTGCTATGTTGGAGCTGGCATTTATTTGCTGAAACCTCTTCC 3422
Db 2941 AGGCATCGTACATAGGTGCTATGTTGGAGCTGGCATTTATTTGCTGAAACCTCTTCC 3000
Qy 3423 ABAAGCAATCAATATGATATGAAATTTGGAGGAGGTACTGGGTGCTCAGTTTCAAAAGAC 3482
Db 3001 ABAAGCAATCAATATGATATGAAATTTGGAGGAGGTACTGGGTGCTCAGTTTCAAAAGAC 3060
Qy 3483 AGATCTGTTTACATTTGCCACAGCAGCTGCTCTTTTCCGGGTAACTTTGGGAAAGTCT 3542
Db 3061 AGATCTGTTTACATTTGCCACAGCAGCTGCTCTTTTCCGGGTAACTTTGGGAAAGTCT 3120
Qy 3543 TTCTCTGAGTTTCAAGTGAATGAAATGGACATTTCTCTCCAGGTCATCACTCAGTCACT 3602
Db 3121 TTCTCTGAGTTTCAAGTGAATGAAATGGACATTTCTCTCCAGGTCATCACTCAGTCACT 3180
Qy 3603 GGTAGGCCAGTCTAATGGCTTAGCATTTAGCTGAATATGTTTATTTACAGAGGAGACAG 3662
Db 3181 GGTAGGCCAGTCTAATGGCTTAGCATTTAGCTGAATATGTTTATTTACAGAGGAGACAG 3240
Qy 3663 GCTTATCTCTGAGTATTTAATTTACTTACCAGATTTATGAGGCTGAAGGTATGGTCGATGGA 3722
Db 3241 GCTTATCTCTGAGTATTTAATTTACTTACCAGATTTATGAGGCTGAAGGTATGGTCGATGGA 3300
Qy 3723 TAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGCAGCAGTGGC 3782
Db 3301 TAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGCAGCAGTGGC 3360
Qy 3783 CTCTAGCTTTTACTCTCTTCTGTAAGAAAAA 3816
Db 3361 CTCTAGCTTTTACTCTCTTCTGTAAGAAAAA 3394

Wed Feb 12 16:08:49 2003

us-09-843-159b-2.rge

Page 33

Search completed: February 11, 2003, 18:26:24
Job time : 9520.64 secs

(RIGE-) RIGEL PHARM INC.

Luo Y, Chan E, Xu X, Huang B;

WPI: 2001-300503/31.

P-PSDB; AAY97749.

Novel recombinant cell cycle polypeptide, tankyrase H useful for inducing or preventing cell proliferation in cells, and for diagnosing, treating or preventing cell cycle associated disorders such as cancer

Claim 13; Fig 2; 63pp; English.

This sequence encodes the Tankyrase homologue isotype 2 (TaHo-2) protein of the invention. The invention also relates to the TaHo-1 protein. The TaHo proteins are useful for inducing or preventing cell proliferation in cells, and in the study or treatment of conditions mediated by the cell cycle proteins, such as to diagnose, treat or prevent cell cycle associated disorders, preferably cancer. The TaHo coding sequences are useful as hybridisation probes, in chromosome and gene mapping and in the generation of anti-sense DNA and RNA. The coding sequences are also useful for the preparation of TaHo, for generating either transgenic animals or knock out animals which, in turn, are useful in a development and screening of therapeutically useful agents, in gene therapy, as vaccine, and for construction of hybridisation probes for mapping the gene which encodes TaHo and for the genetic analysis of individuals with genetic disorders. The TaHo proteins, and their coding sequences are useful in screening assays.

Sequence 3816 BP; 1097 A; 806 C; 949 G; 964 T; 0 other;

Query Match 100.0%; Score 3816; DB 22; Length 3816;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGGTGCTCCCGCGCGCGGCGAGCGGGGCGAGGAGCCAGCGAGGGGCGCGG 60
Db 1 CCGGTGCTCCCGCGCGCGGCGAGCGGGGCGAGGAGCCAGCGAGGGGCGCGG 60
QY 61 TGGGCGCGGCCCATGGAGTGC CGCGGATCCGGTGACACAGGAGGCCAAGCGGCGCGG 120
Db 61 TGGGCGCGGCCCATGGAGTGC CGCGGATCCGGTGACACAGGAGGCCAAGCGGCGCGG 120
QY 121 CCCTGAGCGCTTCTCGGGGGGCGTCCCTCGTCTCGGGGGCGGGGCTCTCTG 180
Db 121 CCCTGAGCGCTTCTCGGGGGGCGTCCCTCGTCTCGGGGGCGGGGCTCTCTG 180
QY 181 TCCGGTTGCTGCGGCTGTGCTGGCTGTGGCGGCGGCCAGGATCATGTGCGGCTCGCGCT 240
Db 181 TCCGGTTGCTGCGGCTGTGCTGGCTGTGGCGGCGGCCAGGATCATGTGCGGCTCGCGCT 240
QY 241 GCGCGCGGGGAGCGGCTCGCGAGCGCGCGCGCGCGGCGGCGGCGGCGGCG 300
Db 241 GCGCGCGGGGAGCGGCTCGCGAGCGCGCGCGCGCGGCGGCGGCGGCGGCG 300
QY 301 GAGAGCTGTTGAGGGTCCCGACAGGGGAGCTGGAAGAGTCAAGAGGCTGGTGACGC 360
Db 301 GAGAGCTGTTGAGGGTCCCGACAGGGGAGCTGGAAGAGTCAAGAGGCTGGTGACGC 360
QY 361 CTGGAAGGTGAACAGCGACGCGGCGGCGAGGAAATCCACCCCGCTGCACTTCGCGG 420
Db 361 CTGGAAGGTGAACAGCGACGCGGCGGCGAGGAAATCCACCCCGCTGCACTTCGCGG 420
QY 421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGAATGTCGAAG 480
Db 421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGAATGTCGAAG 480
QY 481 CACGTGATGATGGGCGCTTATCTCTTCATTAATGTCATGCTCTTTGGTCAATGCTGAAG 540
Db 481 CACGTGATGATGGGCGCTTATCTCTTCATTAATGTCATGCTCTTTGGTCAATGCTGAAG 540
QY 541 TAGTCAATCTCTTTTGGCAGATGGTGACAGACCCCAATGCTCGAGATAATTTGGAATTATA 600

Db 541 TAGTCAATCTCTTTTGGCAGATGGTGACAGACCCCAATGCTCGAGATAATTTGGAATTATA 600
QY 601 CTCCTCTCCATCAAGCTCAATTAAGGAAAGATGATGTTTGCATTTGCTCTTACAGC 660
Db 601 CTCCTCTCCATCAAGCTCAATTAAGGAAAGATGATGTTTGCATTTGCTCTTACAGC 660
QY 661 ATGGAGCTGAGCAACCATCCGAAATACAGATGGAGGAGACAGCATTTGGATTTAGCAGATC 720
Db 661 ATGGAGCTGAGCAACCATCCGAAATACAGATGGAGGAGACAGCATTTGGATTTAGCAGATC 720
QY 721 CATCTGCCAAACAGTCTTACTGGTGAATATAAGAAAGATGAACCTTTAGAAAAGTGCCA 780
Db 721 CATCTGCCAAACAGTCTTACTGGTGAATATAAGAAAGATGAACCTTTAGAAAAGTGCCA 780
QY 781 GGAGTGGCAATCAAGAAATAAGTGGCTCTACTACACCATTAATGTCAACTGCCAGC 840
Db 781 GGAGTGGCAATCAAGAAATAAGTGGCTCTACTACACCATTAATGTCAACTGCCAGC 840
QY 841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGATATAACAGAGTAAAGA 900
Db 841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGATATAACAGAGTAAAGA 900
QY 901 TTGTCAGCTGTTACTGCAACATGGAGCTGATGCCATGCTAAAGATAAAGGTGATCTGG 960
Db 901 TTGTCAGCTGTTACTGCAACATGGAGCTGATGCCATGCTAAAGATAAAGGTGATCTGG 960
QY 961 TACCATTACACAATGCCTGTTCTTATGGTCAATTATGAAGTAACTGAACCTTTGGTCAAGC 1020
Db 961 TACCATTACACAATGCCTGTTCTTATGGTCAATTATGAAGTAACTGAACCTTTGGTCAAGC 1020
QY 1021 ATGGTGCCCTGTAAATGCAATGGACTTTGTGGCAATTCATCTCTTCATGAGGAGCTTT 1080
Db 1021 ATGGTGCCCTGTAAATGCAATGGACTTTGTGGCAATTCATCTCTTCATGAGGAGCTTT 1080
QY 1081 CTAAGAACAGGCTTGAAGTATGTTCTCTCTTAAAGTTATGGTGAGACCCCAACTGTC 1140
Db 1081 CTAAGAACAGGCTTGAAGTATGTTCTCTCTTAAAGTTATGGTGAGACCCCAACTGTC 1140
QY 1141 TCAATTGTCACAATAAAAGTGTATAGACTTTGGCTCCACACACAGTAAAGAAAGAT 1200
Db 1141 TCAATTGTCACAATAAAAGTGTATAGACTTTGGCTCCACACACAGTAAAGAAAGAT 1200
QY 1201 TAGCATATGAATTTAAAGGCCACTCGTCTCGAAGCTGCAGAGAGCTGATGTTACTC 1260
Db 1201 TAGCATATGAATTTAAAGGCCACTCGTCTCGAAGCTGCAGAGAGCTGATGTTACTC 1260
QY 1261 GAATCAAAAACATCTCTCTGGAATGTTCAAGCATCCTCAAAACACATGAAA 1320
Db 1261 GAATCAAAAACATCTCTCTGGAATGTTCAAGCATCCTCAAAACACATGAAA 1320
QY 1321 CAGCATTTGCTGCTGCTGCTATCCATATCCCAAAAAGAAAGCAATATGTGAAGTGT 1380
Db 1321 CAGCATTTGCTGCTGCTGCTATCCATATCCCAAAAAGAAAGCAATATGTGAAGTGT 1380
QY 1381 TCCTAAGAAAGGAGCAACATCAATGAAAAGCTTAAAGAAATTTTGACTCTCTGTCAGC 1440
Db 1381 TCCTAAGAAAGGAGCAACATCAATGAAAAGCTTAAAGAAATTTTGACTCTCTGTCAGC 1440
QY 1441 TGGCATCTCAGAAAGCTCAATATGATGTTGTTGAAGTGTGTTGAAGCATGAAGCAAGG 1500
Db 1441 TGGCATCTCAGAAAGCTCAATATGATGTTGTTGAAGTGTGTTGAAGCATGAAGCAAGG 1500
QY 1501 TTAATGCTCTGGATTAATCTTTGGTTCAGCTTCTTACACAGAGCTGCATATTTGGTTCATC 1560
Db 1501 TTAATGCTCTGGATTAATCTTTGGTTCAGCTTCTTACACAGAGCTGCATATTTGGTTCATC 1560
QY 1561 TACAACCTGCGGCTTACTCTGAGCTATGGGTGATGCTTACATATATATCCCTTCAGC 1620
Db 1561 TACAACCTGCGGCTTACTCTGAGCTATGGGTGATGCTTACATATATATCCCTTCAGC 1620
QY 1621 GCTTTACTCTTTACAGATGGGAAATGAAAATGTTACAGCAACTCTCTCCAGAGGGTATCT 1680
Db 1621 GCTTTACTCTTTACAGATGGGAAATGAAAATGTTACAGCAACTCTCTCCAGAGGGTATCT 1680

Db	1621	GC	TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAAGAGGGTATCT	1680
Qy	1681	CATTAG	TAATTCAGAGGCAGACACACAATTCCTGGAAGCTGCAAAAGCCTGGAGATGTCG	1740
Db	1681	CATTAG	TAATTCAGAGGCAGACACACAATTCCTGGAAGCTGCAAAAGCCTGGAGATGTCG	1740
Qy	1741	AAACTG	TAAAAAATGTTACTGTTTCAGAGTGTCAACTGTCAGAGACATTTGAAGGGCGTC	1800
Db	1741	AAACTG	TAAAAAATGTTACTGTTTCAGAGTGTCAACTGTCAGAGACATTTGAAGGGCGTC	1800
Qy	1801	AGTCTA	CACCACCTTCATTTTGCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGC	1860
Db	1801	AGTCTA	CACCACCTTCATTTTGCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGC	1860
Qy	1861	TACAGC	ATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCTTTGTACCTTTGCACAATG	1920
Db	1861	TACAGC	ATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCTTTGTACCTTTGCACAATG	1920
Qy	1921	CATGTT	CTTATGGACATTAATGAAGTTGCAGAACTTCTTTTAAACATGGAGCAGTAGTTA	1980
Db	1921	CATGTT	CTTATGGACATTAATGAAGTTGCAGAACTTCTTTTAAACATGGAGCAGTAGTTA	1980
Qy	1981	ATGTAG	CTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATG	2040
Db	1981	ATGTAG	CTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATG	2040
Qy	2041	AAATTTG	CAAACTTCCTCCAGCATGGTGCAGACCCCTACCACAAAAACAGGATGGAA	2100
Db	2041	AAATTTG	CAAACTTCCTCCAGCATGGTGCAGACCCCTACCACAAAAACAGGATGGAA	2100
Qy	2101	ATACTC	CTTTGGATCTTTTAAAGATGGAGATPACAGATATTTCAAGATCTGCTTAGGGGAG	2160
Db	2101	ATACTC	CTTTGGATCTTTTAAAGATGGAGATPACAGATATTTCAAGATCTGCTTAGGGGAG	2160
Qy	2161	ATGCAG	CTTGTCTAGATGCTGCCAAGAGGGTTGTTTAGCCAGAGTGAAGAGTTGTCTTT	2220
Db	2161	ATGCAG	CTTGTCTAGATGCTGCCAAGAGGGTTGTTTAGCCAGAGTGAAGAGTTGTCTTT	2220
Qy	2221	CTCCTG	ATATGTAAATTCGCGCATACCCAAGGCAGACATTTCAACACCTTTACATTTAG	2280
Db	2221	CTCCTG	ATATGTAAATTCGCGCGATACCCAAGGCAGACATTTCAACACCTTTACATTTAG	2280
Qy	2281	CAGCTG	GGTTATAATAATTTAGAAGTTGCAGATATTTGTTTACAACACGGAGCTGATGTGA	2340
Db	2281	CAGCTG	GGTTATAATAATTTAGAAGTTGCAGATATTTGTTTACAACACGGAGCTGATGTGA	2340
Qy	2341	ATGCCC	AAGCAAAAGGAGACATTTCCCTTTACATAATGCAGACATCTTAGGGCATGTAG	2400
Db	2341	ATGCCC	AAGCAAAAGGAGACATTTCCCTTTACATAATGCAGACATCTTAGGGCATGTAG	2400
Qy	2401	ATGTAG	CAGCTCTACTTAATAAAGTATATGCAATGTCATGTCCTGTCACGACAAATGGGCTT	2460
Db	2401	ATGTAG	CAGCTCTACTTAATAAAGTATATGCAATGTCATGTCCTGTCACGACAAATGGGCTT	2460
Qy	2461	TCACAC	CTTTGCAGAGCAGCCCAAAAGGGACGAAACACAGCTTTGTGCTTTGTGCTAG	2520
Db	2461	TCACAC	CTTTGCAGAGCAGCCCAAAAGGGACGAAACACAGCTTTGTGCTTTGTGCTAG	2520
Qy	2521	CCCATG	AGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT	2580
Db	2521	CCCATG	AGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT	2580
Qy	2581	CAGCGG	ATGATGTCAGCGCTCTCTTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT	2640
Db	2581	CAGCGG	ATGATGTCAGCGCTCTCTTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT	2640
Qy	2641	GTTACA	AGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAAGGAGCCACTGCAGATGCTCTCT	2700
Db	2641	GTTACA	AGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAAGGAGCCACTGCAGATGCTCTCT	2700
Qy	2701	CTTCAG	GTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAACAACCTATCTG	2760
Db	2701	CTTCAG	GTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAACAACCTATCTG	2760

RESULT 2

AAC85294
 ID AAC85294 standard; cDNA; 6019 BP.
 XX AC AAC85294;
 XX DT 29-MAR-2001 (first entry)
 XX DE Human SPANK cDNA.
 XX KW SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
 KW poly adenosine diphosphate-ribose polymerase; catalytic domain;
 KW ANK; ankyrin repeat; cytosol; insulin-responsive aminopeptidase;
 KW IRAP; GLUT4; adipocyte; insulin signalling pathway; hyperlipidaemia;
 KW glucose intolerance; atherosclerotic disease; atherosclerosis;
 KW obesity; cardiac insufficiency; coronary insufficiency; stroke;
 KW high blood pressure; non-insulin dependent diabetes; hypertension;
 KW hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy; ds.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX CDS 252...3735
 XX FT /*tag= a
 XX FT /product= "Human SPANK"
 XX FT /transl_except= (pos:768..878,aa:173..209)
 XX FT /transl_except= (pos:1230..1337,aa:327..362)
 XX FT /transl_except= (pos:1374..1445,aa:375..398)
 XX FT /transl_except= (pos:1695..1823,aa:482..524)
 XX FT /transl_except= (pos:2172..2282,aa:641..677)
 XX PN WO200077225-A1.
 XX PD 21-DEC-2000.
 XX PF 09-JUN-2000; 2000WO-US15926.
 XX PR 11-JUN-1999; 99US-0138957.
 XX PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
 XX PA (GEO) GEN HOSPITAL CORP.
 XX PI Chi N, Lodish HF;
 XX WPI; 2001-091404/10.
 XX DR P-PSDB; AAB47022.
 XX PT New insulin signalling protein SPANK, useful for reducing body mass,
 PT glucose intolerance or insulin resistance and for preventing or
 PT treating obesity-related and muscle-related diseases -
 XX PS Claim 3; Fig 4; 65pp; English.
 XX CC This sequence represents the human SPANK cDNA. The SPANK protein
 CC comprises 3 domains:
 CC (a) a SAM (sterile alpha motif) domain;
 CC (b) a PARP (poly adenosine diphosphate-ribose polymerase) catalytic
 CC domain; and
 CC (c) an ANK domain composed of ankyrin repeats.
 CC SPANK is a cytosolic protein which can poly(ADP-riboseyl)ate itself.
 CC SPANK binds insulin-responsive aminopeptidase (IRAP) and modulates
 CC translocation of GLUT4 in the perinuclear region of adipocytes. It
 CC is an effector in the insulin signalling pathway in eukaryotic cells.
 CC SPANK is useful for reducing body mass, reducing glucose
 CC intolerance or insulin resistance, for preventing or treating
 CC obesity-related diseases or disorders, such as obesity, cardiac
 CC insufficiency, coronary insufficiency, stroke, hypertension,
 CC atherosclerotic disease, hyperlipidaemia, hyperuricemia and Syndrome X and is
 CC dependent diabetes, hyperlipidaemia, hyperuricemia and Syndrome X and is
 CC also useful for preventing or treating muscle-related diseases or
 CC disorders, such as muscular dystrophy, muscle atrophy and muscle
 CC fatigue. Antibodies immunospecific for SPANK are useful for detecting

CC the presence of SPANK polypeptide in a biological sample.

XX Sequence 6019 BP; 1732 A; 1173 C; 1341 G; 1773 T; 0 other;

SQ

Query Match 99.4%; Score 3793; DB 22; Length 6019;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3807; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 CGCGCTGCTCCCGCCCGCGGCGAGCCGGGGGAGGAGCCAGCCAGGAGGCGCGCG 60

DB 29 CGCGCTGCTCCCGCCCGCGGCGAGCCGGGGGAGGAGCCAGCCAGGAGGCGCGCG 88

QY 61 TGGGCGCGCCCGATGGGACTGGCCCGATCCCGGTGACAGCAGGAGCCAAAGCGCCCGGG 120

DB 89 TGGGCGCGG-CGATGGGACTGGCCCGATCCCGGTGACAGCAGGAGCCAAAGCGCCCGGG 147

QY 121 CCCTGAGCGCGTCTTCTCCGGGGGGGCTCGCCCTCTCTGCTCGCGGGCGCGGGCTCTCTGC 180

DB 148 CCTGAGCGCGTCTTCTCCGGGGGGGCTCGCCCTCTCTGCTCGCGGGCGCGGGCTCTCTGC 207

QY 181 TCGGTTGCTGCGCTGTTGCTGGCTGTGGCGCGCGCCAGGATCATGTGCGGTGCGCGCT 240

DB 208 TCGGTTGCTGCGCTGTTGCTGGCTGTGGCGCGCGCCAGGATCATGTGCGGTGCGCGCT 267

QY 241 GCGCGCGGGGAGCGGCTGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300

DB 268 GCGCGCGGGGAGCGGCTGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 327

QY 301 GAGAGCTGTTCCAGGCGTCCCGCAACGCGGAGCGTGAAGCAGTCAAGAGCGTGGTGCACGC 360

DB 328 GAGAGCTGTTCCAGGCGTCCCGCAACGCGGAGCGTGAAGCAGTCAAGAGCGTGGTGCACGC 387

QY 361 CTGAGAAGGTGAACAGCGCGCACCGCGCGCGGAGGAGAAATCCACCCGCTGCACCTTCGCGCG 420

DB 388 CTGAGAAGGTGAACAGCGCGCACCGCGCGCGGAGGAGAAATCCACCCGCTGCACCTTCGCGCG 447

QY 421 CAGGTTTGGCGGGAAGAGCTAGTTGAATATTTGCTTCAGAAATGTGCAAAATGTCAAG 480

DB 448 CAGGTTTGGCGGGAAGAGCTAGTTGAATATTTGCTTCAGAAATGTGCAAAATGTCAAG 507

QY 481 CACGTGATGATGGGGCTTATTCCTCTTCATATCATGCTCTTTTGGTTCATGCTGAAG 540

DB 508 CACGTGATGATGGGGCTTATTCCTCTTCATATCATGCTCTTTTGGTTCATGCTGAAG 567

QY 541 TAGTCAATCTCCTTTTGGACATGTGTCAGACCCCAATGCTCGAGATAATTTGGAATATTA 600

DB 568 TAGTCAATCTCCTTTTGGACATGTGTCAGACCCCAATGCTCGAGATAATTTGGAATATTA 627

QY 601 CTCCTCTCATGAAGCTGCAATTAAGGAAGATTGATGTTTGCATGTTGCTGTACAGC 660

DB 628 CTCCTCTCATGAAGCTGCAATTAAGGAAGATTGATGTTTGCATGTTGCTGTACAGC 687

QY 661 ATGGAGCTGAGCAACCATCCGAAATACAGATGGAAGAGCAGCATTTGGATTAGCAGATC 720

DB 688 ATGGAGCTGAGCAACCATCCGAAATACAGATGGAAGAGCAGCATTTGGATTAGCAGATC 747

QY 721 CATCTGCCAAAGCAGTGTCTTACTGTTGAATATAAGAAAGATGAACCTCTTAGAAAGTGCCA 780

DB 748 CATCTGCCAAAGCAGTGTCTTACTGTTGAATATAAGAAAGATGAACCTCTTAGAAAGTGCCA 807

QY 781 GGAGTGGCAATGAAGAAATAATGCTGCTACTACACCATTAATGTCACCTGCCACG 840

DB 808 GGAGTGGCAATGAAGAAATAATGCTGCTACTACACCATTAATGTCACCTGCCACG 867

QY 841 CAAGTGATGGCAAGAGTCAACCTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGA 900

DB 868 CAAGTGATGGCAAGAGTCAACCTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGA 927

QY 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGCTCATGCTAAAGATAAAGGTGATCTGG 960

DB 928 TTGTACAGCTGTTACTGCAACATGGAGCTGATGCTCATGCTAAAGATAAAGGTGATCTGG 987

QY 961 TACCATTACAAATGCTCTTCTTATGGTCAATTATGAAGTAACTGAACCTTTTGGTCAAGC 1020

Db 988 |||||TACCAATACCAATGCGCTGTTCTTATGTCATATTATGAAGTAACTGAACTTTTGGTCAAGC 1047
QY 1021 ATGGTGCCTGTGTAATGCAATGGACTTGGCAATTGCGCAATTTCACCTCTCTTTCATGAGCGAGCTT 1080
Db 1048 ATGGTGCCTGTGTAATGCAATGGACTTGGCAATTGCGCAATTTCACCTCTCTTTCATGAGCGAGCTT 1107
QY 1081 CTAAAGAACAGGGTTGAAGTATGTTCTCTCTTCTTAAAGTTATGTTGTCAGACCCCAACACTGC 1140
Db 1108 CTAAAGAACAGGGTTGAAGTATGTTCTCTCTTCTTAAAGTTATGTTGTCAGACCCCAACACTGC 1167
QY 1141 TCAATTTGTCACAATAAAGTGTCTATAGACTTGGCTGCCACACACACAGTTTAAAGAAAGAT 1200
Db 1168 TCAATTTGTCACAATAAAGTGTCTATAGACTTGGCTGCCACACACACAGTTTAAAGAAAGAT 1227
QY 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGTGTGCAGAGAGCTGATGTTACTC 1260
Db 1228 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGTGTGCAGAGAGCTGATGTTACTC 1287
QY 1261 GAATCAAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGCATCCTCAAAACACATGAAA 1320
Db 1288 GAATCAAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGCATCCTCAAAACACATGAAA 1347
QY 1321 CAGCATTTGCAATTTGCTGCTGCATCTCCATATCCAAAAGAAAGCAAAATATGTAAGTGT 1380
Db 1348 CAGCATTTGCAATTTGCTGCTGCATCTCCATATCCAAAAGAAAGCAAAATATGTAAGTGT 1407
QY 1381 TGCTAAGAAAAGGAGCAAAACATCAATGAAAGAACTAAAGAAATTTTGACTCCTCTGCAAG 1440
Db 1408 TGCTAAGAAAAGGAGCAAAACATCAATGAAAGAACTAAAGAAATTTTGACTCCTCTGCAAG 1467
QY 1441 TGCAATCTGAGAAAAGCTCAATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGG 1500
Db 1468 TGCAATCTGAGAAAAGCTCAATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGG 1527
QY 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTTACACAGAGCTGCATATTTGGTGCATC 1560
Db 1528 TTAATGCTCTGGATAATCTTGGTCAGACTTCTTACACAGAGCTGCATATTTGGTGCATC 1587
QY 1561 TACAACCTGCCGCCACTCTCTGAGCTATGGGTGTGATCCTCAACATTATATCCCTTCAGG 1620
Db 1588 TACAACCTGCCGCCACTCTCTGAGCTATGGGTGTGATCCTCAACATTATATCCCTTCAGG 1647
QY 1621 GCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCAAAGAGGGTATCT 1680
Db 1648 GCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCAAAGAGGGTATCT 1707
QY 1681 CATTAGGTAATTCAGAGGCAGACAGACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCG 1740
Db 1708 CATTAGGTAATTCAGAGGCAGACAGACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCG 1767
QY 1741 AAAGTGTAAAAAAGCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db 1768 AAAGTGTAAAAAAGCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1827
QY 1801 AGTCTACACCACCTTCATTTTGCAGCTGGGTATTAACAGAGTGTCCGGTGGGAATATCTGC 1860
Db 1828 AGTCTACACCACCTTCATTTTGCAGCTGGGTATTAACAGAGTGTCCGGTGGGAATATCTGC 1887
QY 1861 TACAGCATGGAGCTGATGTCATGCTTAAGATATAAGATAAGAGGCCCTTGACCTTTGCACAATG 1920
Db 1888 TACAGCATGGAGCTGATGTCATGCTTAAGATATAAGATAAGAGGCCCTTGACCTTTGCACAATG 1947
QY 1921 CATGTTCTTTATGACATTTATGAAGTTGCAAGCTTCTGTTAAACATGGAGCAGTAGTTA 1980
Db 1948 CATGTTCTTTATGACATTTATGAAGTTGCAAGCTTCTGTTAAACATGGAGCAGTAGTTA 2007
QY 1981 ATGTAGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGAGCAAAAGGAAATATG 2040
Db 2008 ATGTAGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGAGCAAAAGGAAATATG 2067
QY 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTCCAGACCTTACCAAAAAAACAGGATGGAA 2100
Db |||||TACCAATACCAATGCGCTGTTCTTATGTCATATTATGAAGTAACTGAACTTTTGGTCAAGC 1047

Db 2068 AAATTTGCAAACTTCTGCTCCAGCATGGTCCAGACCTTACAAAAAACAAGGATGGAA 2127
QY 2101 ATACTCCTTTGGACTCTTGTAAAGATGAGATACAGATATTTCAGACTCTGCTTAGGGAG 2160
Db 2128 ATACTCCTTTGGACTCTTGTAAAGATGAGATACAGATATTTCAGACTCTGCTTAGGGAG 2187
QY 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAGGGTCTTTTAGCCAGAGTGAAGAGTTGTCTT 2220
Db 2188 ATGCAGCTTTGCTAGATGCTGCCAAGAGGGTCTTTTAGCCAGAGTGAAGAGTTGTCTT 2247
QY 2221 CTCTGTATTAATGTAATTTGCCGCGATACCCAAAGCAGACATTTCAACACCTTTACATTTAG 2280
Db 2248 CTCTGTATTAATGTAATTTGCCGCGATACCCAAAGCAGACATTTCAACACCTTTACATTTAG 2307
QY 2281 CAGCTGCTTTAATAATAATTTAGAACTTGCAGAGTATTGTTTACAACACGAGCTGATGTGA 2340
Db 2308 CAGCTGCTTTAATAATAATTTAGAACTTGCAGAGTATTGTTTACAACACGAGCTGATGTGA 2367
QY 2341 ATGCCAAGCAAAAGGAGGACTTATTCTCTTTACATAATGTCAGCATCTTTACGGGATGTAG 2400
Db 2368 ATGCCAAGCAAAAGGAGGACTTATTCTCTTTACATAATGTCAGCATCTTTACGGGATGTAG 2427
QY 2401 ATGTAGCAGCTCTACTAATAAAGTATTAATGATGTGTCATGTCCACGACAAAATGGGCTT 2460
Db 2428 ATGTAGCAGCTCTACTAATAAAGTATTAATGATGTGTCATGTCCACGACAAAATGGGCTT 2487
QY 2461 TCACACCTTTGCGAAGCAGAGCCCAAAAGGACCAACACAGCTTTGTCTTGTGTGTAG 2520
Db 2488 TCACACCTTTGCGAAGCAGAGCCCAAAAGGACCAACACAGCTTTGTCTTGTGTGTAG 2547
QY 2521 CCCATGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTT 2580
Db 2548 CCCATGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTT 2607
QY 2581 CAGCGGATGATGTGACGGCTCTTCTGACAGAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db 2608 CAGCGGATGATGTGACGGCTCTTCTGACAGAGCCATGCCCCCATCTGCTCTGCCCTCTT 2667
QY 2641 GTTACAAGCTCAAGTCTCAATGCTGTGAGAAGCCAGGAGCCACTGTCAGATGCTCTCT 2700
Db 2668 GTTACAAGCTCAAGTCTCAATGCTGTGAGAAGCCAGGAGCCACTGTCAGATGCTCTCT 2727
QY 2701 CTTCAGCTTCCATCTAGCCCCATCAAGCCTTTCTGACGACAGCCAGCTTTTGACAACCTTATCTG 2760
Db 2728 CTTCAGCTTCCATCTAGCCCCATCAAGCCTTTCTGACGACAGCCAGCTTTTGACAACCTTATCTG 2787
QY 2761 GGAGTTTTTTCAGAACTCTCTTTCAGTAGTTAGTTCAAGTGGAAACAGAGGCTTCCAGTT 2820
Db 2788 GGAGTTTTTTCAGAACTCTCTTTCAGTAGTTAGTTCAAGTGGAAACAGAGGCTTCCAGTT 2847
QY 2821 TGGAGAAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db 2848 TGGAGAAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2907
QY 2881 GACTTTGAGCACCTTAATGGATATATTTGAGAGAGAACAGATCAGCTTTGGATGTATTAGTTG 2940
Db 2908 GACTTTGAGCACCTTAATGGATATATTTGAGAGAGAACAGATCAGCTTTGGATGTATTAGTTG 2967
QY 2941 AGATGGGCGACACAGGAGCTGAAGGAGATTGGAATCAATGCTTTATGGACATAGGCACAAAC 3000
Db 2968 AGATGGGCGACACAGGAGCTGAAGGAGATTGGAATCAATGCTTTATGGACATAGGCACAAAC 3027
QY 3001 TAATTAAGGAGCTGAGAGACTTATCTCCGACAAACAGGCTTTAACCCATATTAACTT 3060
Db 3028 TAATTAAGGAGCTGAGAGACTTATCTCCGACAAACAGGCTTTAACCCATATTAACTT 3087
QY 3061 TGAACACCTCTGTTAGTGGAAACAAATTTCTTATAGATCTGTCTCTCTGATGATAAAGAGTTTC 3120
Db 3088 TGAACACCTCTGTTAGTGGAAACAAATTTCTTATAGATCTGTCTCTCTGATGATAAAGAGTTTC 3147
QY 3121 AGTCTGTGGAGGAGAGATGCAAAAGTACAGTTCCAGAGCAGACAGATGGAGGTCATGCAG 3180
Db 3148 AGTCTGTGGAGGAGAGATGCAAAAGTACAGTTCCAGAGCAGACAGATGGAGGTCATGCAG 3207

QY 3181 GTGGAATCTTCAACAGATACAAATATTTCTCAAGATTGAGAAGTTTGTAAACAAGAACTAT 3240
DB 3208 GTGGAATCTTCAACAGATACAAATATTTCTCAAGATTGAGAAGTTTGTAAACAAGAACTAT 3267
QY 3241 GGGAAAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCAATGCCAATG 3300
DB 3268 GGGAAAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCAATGCCAATG 3327
QY 3301 AACGAATGTAATTTATGGGTCTCTCTTTGTGAATGCAATTTATCCACAAGGCTTTGATG 3360
DB 3328 AACGAATGTAATTTATGGGTCTCTCTTTGTGAATGCAATTTATCCACAAGGCTTTGATG 3387
QY 3361 AAAGGATCGGTACATAGTGTGATGTTTGGAGCTGGCAATTTATTTTGTGAAAACCTCTT 3420
DB 3388 AAAGGATCGGTACATAGTGTGATGTTTGGAGCTGGCAATTTATTTGTGAAAACCTCTT 3447
QY 3421 CCAAAAGCAATCAATATGATATGTAATTTGGAGGAGTACTGGGTGTCAGTTCAACAAG 3480
DB 3448 CCAAAAGCAATCAATATGATATGTAATTTGGAGGAGTACTGGGTGTCAGTTCAACAAG 3507
QY 3481 ACAGATCTTTGATACATTTCCACAGCAGCTGCTCTTTTCCGCGGTAACTTTGGGAAAGT 3540
DB 3508 ACAGATCTTTGATACATTTCCACAGCAGCTGCTCTTTTCCGCGGTAACTTTGGGAAAGT 3567
QY 3541 CTTTCTCGAGTTCAGTGCATGAAATGGCACATTTCTCTCCAGGTCACTACTCAGTCA 3600
DB 3568 CTTTCTCGAGTTCAGTGCATGAAATGGCACATTTCTCTCCAGGTCACTACTCAGTCA 3627
QY 3601 CTGGTAGGCCAGTGAATAGCCCTAGCAATAGCTGAATATGTTATTTACAGAGGAGAAC 3660
DB 3628 CTGGTAGGCCAGTGAATAGCCCTAGCAATAGCTGAATATGTTATTTACAGAGGAGAAC 3687
QY 3661 AGGCTTATCTCAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3720
DB 3688 AGGCTTATCTCAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3747
QY 3721 GATAAATAGTATTTTAAAGAACTAATTCCTACTGAACCTTAAATATCAAAAGCAGCAGTG 3780
DB 3748 GATAAATAGTATTTTAAAGAACTAATTCCTACTGAACCTTAAATATCAAAAGCAGCAGTG 3807
QY 3781 GCCTCTACGTTTTTACTCTTTGCTGAAAAAAA 3813
DB 3808 GCCTCTACGTTTTTACTCTTTGCTGAAAAAAA 3840

RESULT 3
AAC66825
ID AAC66825 standard; cDNA; 4275 BP.
XX AC AAC66825;
XX AC

XX 27-FEB-2001 (first entry)

XX Human tankyrase II coding sequence SEQ ID NO: 5.
DE Human; tankyrase II; telomere length; signal transduction; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
FH 284..3784
CDS /*tag= a
FT /product= "tankyrase II"
FT /transl_except= (pos:655..660,aa:"IyrThr")
FT

XX W0200061813-A1.

XX 19-OCT-2000.

XX 10-APR-2000; 2000WO-US09558.

XX 09-APR-1999; 99US-0128577.

PR 13-APR-1999; 99US-0129123.
XX (GERO-) GERON CORP.
XX PI Morin GB, Funk WD, Piatyszek MA;
XX WPI; 2000-679503/66.
DR P-PSDB: AAB27211.
XX Novel mammalian Tankyrase II polypeptide and the polynucleotide
PT encoding the polypeptide useful for modulating or maintaining telomere
PT length, replicative capacity, apoptosis, chromosome packing or gene
PT expression -
XX Claim 1; Fig 4; 52pp; English.
XX The present sequence is a version of the human tankyrase II coding
CC sequence. Its protein is thought to be involved in signal transduction in
CC the cell, and to have binding activity for other telomere-associated
CC proteins. It is possible that it plays a role in the regulation of
CC telomere length, thus affecting the replicative ability of the cell. The
CC protein is useful for ribosylating target proteins, for determining
CC tankyrase II binding activity in a sample, and for modulating telomere
CC length in a cell.
XX Sequence 4275 BP; 1247 A; 889 C; 1025 G; 1114 T; 0 other;
Query Match 99.4%; Score 3791.4; DB 21; Length 4275;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 1 CGCGTGTCTCCCGCCCGCGCGGAGCCAGCGGGGAGAGCCAGCGAGGGCGCGCG 60
DB 61 CGCGTGTCTCCCGCCCGCGCGGAGCCAGCGGGGAGAGCCAGCGAGGGCGCGCG 120
QY 61 TGGGCGCGCGCCATGAGACTGCGCGGATCCCGGTGACACAGAGGAGCCAAAGCGCCCGG 120
DB 121 TGGGCGCGCG-CCATGAGACTGCGCGGATCCCGGTGACACAGAGGAGCCAAAGCGCCCGG 179
QY 121 CCCTGAGCGCTTCTCCGCGGGGCGCTCGCCCTCTCTGCTCGCGGGCGCGGGCTCTCG 180
DB 180 CCCTGAGCGCTTCTCTCCGCGGGGCGCTCGCCCTCTCTGCTCGCGGGCGCGGGCTCTCG 239
QY 181 TCCGCTTGTGCGCTGTTGCTGTGTGCGCGCGCGCCAGGATCATGTGCGGTGCGCGCT 240
DB 240 TCCGCTTGTGCGCTGTTGCTGTGTGCGCGCGCGCCAGGATCATGTGCGGTGCGCGCT 299
QY 241 GCGCGCGCGGGGAGCGCGCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 300 GCGCGCGCGGGGAGCGCGCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 359
QY 301 GAGAGCTGTTGAGGCGTCCCGCAACGGGAGCTGGAACGAGTCAAGAGGCTGCTGACGC 360
DB 360 GAGAGCTGTTGAGGCGTCCCGCAACGGGAGCTGGAACGAGTCAAGAGGCTGCTGACGC 419
QY 361 CTGAGAAGGTGAACAGCGCGACACCGCGGGGAGGAAATCCACCCCGCTGCACTTCGCGG 420
DB 420 CTGAGAAGGTGAACAGCGCGACACCGCGGGGAGGAAATCCACCCCGCTGCACTTCGCGG 479
QY 421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTCAGAAATGTCAAATGTCGAAG 480
DB 480 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTCAGAAATGTCAAATGTCGAAG 539
QY 481 CAGGTGATGAGGCGCTTATTCCTCTTCAATAATGCATGCTCTTTTGTGCTGCTGAAG 540
DB 540 CAGGTGATGAGGCGCTTATTCCTCTTCAATAATGCATGCTCTTTTGTGCTGCTGAAG 599
QY 541 TAGTCAATCTCTTTTTCGACATGTTGTCAGACCCCAATGCTCGAGATAATTTGGAATATA 600
DB 600 TAGTCAATCTCTTTTTCGACATGTTGTCAGACCCCAATGCTCGAGATAATTTGGAATATA 659
QY 601 CTCCTCTCATGAGCTGCAATTAAGGAAAGATTGATGTTTGCATGCTGCTTACAGC 660
DB 601 CTCCTCTCATGAGCTGCAATTAAGGAAAGATTGATGTTTGCATGCTGCTTACAGC 660

Db 660 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGGCAITGTGCTGTACAGC 719
Qy 661 ATGGAGCTGAGCAACCACTCCGAAATACAGATGGAAGACAGCACTTGATTTAGCAGATC 720
Db 720 ATGGAGCTGAGCAACCACTCCGAAATACAGATGGAAGACAGCACTTGATTTAGCAGATC 779
Qy 721 CATGTGCCAAAGCACTGCTTACTGTGTAATATPAGAAAGATGAACCTCTTAAAGAGTGCCA 780
Db 780 CATCTGCCAAAGCACTGCTTACTGTGTAATATPAGAAAGATGAACCTCTTAAAGAGTGCCA 839
Qy 781 CGAGTGGCAATGAAGAAATATGAGCTCTACTCACACCATTAAATGTCACCTGCCACG 840
Db 840 GGAGTGGCAATGAAGAAATATGAGCTCTACTCACACCATTAAATGTCACCTGCCACG 899
Qy 841 CAAGTATGGCAAGCACTCCCACTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db 900 CAAGTATGGCAAGCACTCCCACTACATTTGGCAGCAGGATATAACAGAGTAAAGA 959
Qy 901 TTGTACAGCTGTTACTGCAACATGAGCTGATGTCCTATGAAGATAAAGTGATCTGG 960
Db 960 TTGTACAGCTGTTACTGCAACATGAGCTGATGTCCTATGAAGATAAAGTGATCTGG 1019
Qy 961 TACCATTACACAATGCTGTTCTTATGCTCAATTATGAAGTAACTCAACTTTTGGTCAAGC 1020
Db 1020 TACCATTACACAATGCTGTTCTTATGCTCAATTATGAAGTAACTCAACTTTTGGTCAAGC 1079
Qy 1021 ATGGTGCCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTCATGAGGAGCTT 1080
Db 1080 ATGGTGCCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTCATGAGGAGCTT 1139
Qy 1081 CTAAGAACAGGTTGAAGTATGTTCTCTCTCTTAAAGTATGTTGGTGCAGACCCCAACACTGC 1140
Db 1140 CTAAGAACAGGTTGAAGTATGTTCTCTCTCTTAAAGTATGTTGGTGCAGACCCCAACACTGC 1199
Qy 1141 TCAATGTCACAATRAAGTGCTATAGACTTGGCTCCACACCCAGTTAAAGAAAGAT 1200
Db 1200 TCAATGTCACAATRAAGTGCTATAGACTTGGCTCCACACCCAGTTAAAGAAAGAT 1259
Qy 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTCAAGCTGCAGAGCAAGCTGATGTTACTC 1260
Db 1260 TAGCATATGAATTTAAAGGCCACTCGTTGCTCAAGCTGCAGAGCAAGCTGATGTTACTC 1319
Qy 1261 GAATCAAAAACATCTCTCTCTGGAATGGTGAAATTTCAAGCATCCTCAACACATGAAA 1320
Db 1320 GAATCAAAAACATCTCTCTCTGGAATGGTGAAATTTCAAGCATCCTCAACACATGAAA 1379
Qy 1321 CAGCATTTGCATTTGCTGCTGCTATCCATATCCCAAGAAAGCAAAATATGTGAACGT 1380
Db 1380 CAGCATTTGCATTTGCTGCTGCTATCCATATCCCAAGAAAGCAAAATATGTGAACGT 1439
Qy 1381 TGCTAAGAAAGGAGCAACATCAATGAAAGACTTAAAGAAATTCCTGACTCTCTGCAAG 1440
Db 1440 TGCTAAGAAAGGAGCAACATCAATGAAAGACTTAAAGAAATTCCTGACTCTCTGCAAG 1499
Qy 1441 TGGCATCTGAGAAAGCTCAATATGATGTTGTTGAAAGTGGTGAACATGAAGCAAAAG 1500
Db 1500 TGGCATCTGAGAAAGCTCAATATGATGTTGTTGAAAGTGGTGAACATGAAGCAAAAG 1559
Qy 1501 TTAATGCTCTGGATATCTTGGCTGAGACTTCTACACAGAGCTGCATATGTGGTCAATC 1560
Db 1560 TTAATGCTCTGGATATCTTGGCTGAGACTTCTACACAGAGCTGCATATGTGGTCAATC 1619
Qy 1561 TACAAACCTGCGCCCTACTCTCTGAGCTATGGGTGATCCTAACATATATCCCTTCAGG 1620
Db 1620 TACAAACCTGCGCCCTACTCTCTGAGCTATGGGTGATCCTAACATATATCCCTTCAGG 1679
Qy 1621 GCTTTACTGCTTACAGATGGGAAATGAAATGTACAGCAACTCTCTCCAAAGAGGATATCT 1680
Db 1680 GCTTTACTGCTTACAGATGGGAAATGAAATGTACAGCAACTCTCTCCAAAGAGGATATCT 1739
Qy 1681 CATTAGGTAAATCAGAGGAGACAGACAAATTCCTGGAAGCTGCAAGGCTGGAGATGCG 1740
Db 1740 CATTAGGTAAATCAGAGGAGACAGACAAATTCCTGGAAGCTGCAAGGCTGGAGATGCG 1799

Qy 1741 AAACGTAAAAAACTCTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTAAAGGCGTC 1800
Db 1800 AAACGTAAAAAACTCTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTAAAGGCGTC 1859
Qy 1801 AGCTACACCACTTCATTTTGCAGCTGGGTATACAGAGTGTCCGGTGGGAATATCTGC 1860
Db 1860 AGCTACACCACTTCATTTTGCAGCTGGGTATACAGAGTGTCCGGTGGGAATATCTGC 1919
Qy 1861 TACAGCATGAGCTGATGTCATGCTAAAGATAAGAGAGCCTTGTACCTTTGCACAATG 1920
Db 1920 TACAGCATGAGCTGATGTCATGCTAAAGATAAGAGAGCCTTGTACCTTTGCACAATG 1979
Qy 1921 CATGCTCTTATGGACATTTAAGTTCAGAACTTCTTGTAAACATGGAGCAGTAGTTA 1980
Db 1980 CATGCTCTTATGGACATTTAAGTTCAGAACTTCTTGTAAACATGGAGCAGTAGTTA 2039
Qy 1981 ATGTAGCTGATTTATGAAATTTTACACTTTTACATGAAGCAGCAGCAAAAGGAAATATG 2040
Db 2040 ATGTAGCTGATTTATGAAATTTTACACTTTTACATGAAGCAGCAGCAAAAGGAAATATG 2099
Qy 2041 AAATTTGCAAACTTCTCTCCAGCATGGTCAGACCCTACCAAAAAAACACAGGATGAA 2100
Db 2100 AAATTTGCAAACTTCTCTCCAGCATGGTCAGACCCTACCAAAAAAACACAGGATGAA 2159
Qy 2101 ATACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGAG 2160
Db 2160 ATACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGAG 2219
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAGTTGCTT 2220
Db 2220 ATGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAGTTGCTT 2279
Qy 2221 CTCTGTATTAATGTAATTTGCCGATACCAAGCAGACATTCACACCTTTACATTTAG 2280
Db 2280 CTCTGTATTAATGTAATTTGCCGATACCAAGCAGACATTCACACCTTTACATTTAG 2339
Qy 2281 CAGCTGGTTAATAATTTAGAACTTGCAGAGTATTTGTTACACACGAGCTGATGTGA 2340
Db 2340 CAGCTGGTTAATAATTTAGAACTTGCAGAGTATTTGTTACACACGAGCTGATGTGA 2399
Qy 2341 ATGCCCAAGACAAAGGAGGACTTATCTTTTACATATGAGCATCTTACGGGCATGTAG 2400
Db 2400 ATGCCCAAGACAAAGGAGGACTTATCTTTTACATATGAGCATCTTACGGGCATGTAG 2459
Qy 2401 ATGTAGCAGCTCTACTTAATAAGTATTAATGATGTCAATGCCACGACAAAATGGGCTT 2460
Db 2460 ATGTAGCAGCTCTACTTAATAAGTATTAATGATGTCAATGCCACGACAAAATGGGCTT 2519
Qy 2461 TCACACCTTTGCGAAGCAGCCCCAAAGGACACACAGCTTTGTGCTTTGTGCTAG 2520
Db 2520 TCACACCTTTGCGAAGCAGCCCCAAAGGACACACAGCTTTGTGCTTTGTGCTAG 2579
Qy 2521 CCAATGAGCTGAGCCGCACTCTTAAATAATCAGGAAGACAAACACCTTTAGATTTAGTTT 2580
Db 2580 CCAATGAGCTGAGCCGCACTCTTAAATAATCAGGAAGACAAACACCTTTAGATTTAGTTT 2639
Qy 2581 CAGCGGATGATGTGAGGCTCTTCTGACAGCAGCATGCCCCCATCTGCTGCGCTCTT 2640
Db 2640 CAGCGGATGATGTGAGGCTCTTCTGACAGCAGCATGCCCCCATCTGCTGCGCTCTT 2699
Qy 2641 GTTACAAGCCTTCAAGTGTCTCAATGTTGAGAAGCCAGGAGCCACCTGAGATGCTCTCT 2700
Db 2700 GTTACAAGCCTTCAAGTGTCTCAATGTTGAGAAGCCAGGAGCCACCTGAGATGCTCTCT 2759
Qy 2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGAGCAGCAGCAGCTCTTGACAACTTATCTG 2760
Db 2760 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGAGCAGCAGCAGCTCTTGACAACTTATCTG 2819
Qy 2761 GGAGTTTTTCAGAACTGTCTTCACTAGTTAGTTCAAGTGGAAACAGAGGCTTCCAGTT 2820
Db 2820 GGAGTTTTTCAGAACTGTCTTCACTAGTTAGTTCAAGTGGAAACAGAGGCTTCCAGTT 2879

```
QY 2821 TGGAGAAAAGAGGAGTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db 2880 TGGAGAAAAGAGGAGTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2939
QY 2881 GACTTGAGCACCTAAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTAATAGTTG 2940
Db 2940 GACTTGAGCACCTAAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTAATAGTTG 2999
QY 2941 AGATGGGGCAAGAGGAGCTGAAGCAGATGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db 3000 AGATGGGGCAAGAGGAGCTGAAGCAGATGGAATCAATGCTTATGGACATAGGCACAAAC 3059
QY 3001 TAATTAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGCTTTAACCCATATTTAACTT 3060
Db 3060 TAATTAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGCTTTAACCCATATTTAACTT 3119
QY 3061 TGAACACCTCTGTAGTGAACAAATTTCTTATAGATCTGTCTCTGATGATAAAGAGTTTC 3120
Db 3120 TGAACACCTCTGTAGTGAACAAATTTCTTATAGATCTGTCTCTGATGATAAAGAGTTTC 3179
QY 3121 AGTCTGTGGAGGAAGAGATGCAAAATACAGTTGAGAGCACAGAGATGGAGGTCAATGCAG 3180
Db 3180 AGTCTGTGGAGGAAGAGATGCAAAATACAGTTGAGAGCACAGAGATGGAGGTCAATGCAG 3239
QY 3181 GTGGAATCTTCAACAGATACAATATTTCTCAAGATTCAGAAGTTTCTAAACAAGAACTAT 3240
Db 3240 GTGGAATCTTCAACAGATACAATATTTCTCAAGATTCAGAAGTTTCTAAACAAGAACTAT 3299
QY 3241 GGAAGAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCACTGCAAAATG 3300
Db 3300 GGAAGAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCACTGCAAAATG 3359
QY 3301 AACGAATGCTATTTTCATGGTCTCCTTTTGTGAATCAATATPCCACAAGGCTTTTGATG 3360
Db 3360 AACGAATGCTATTTTCATGGTCTCCTTTTGTGAATCAATATPCCACAAGGCTTTTGATG 3419
QY 3361 AAAGGATCGGTACATAGTGGTATGTTGGAGCTGGCAITATTTTGTGAAACTCTT 3420
Db 3420 AAAGGATCGGTACATAGTGGTATGTTGGAGCTGGCAITATTTTGTGAAACTCTT 3479
QY 3421 CCAAAAGCAATCAATATGTATATGGAATGGAGGAGTACTGGTGTCCAGTTCACAAAG 3480
Db 3480 CCAAAAGCAATCAATATGTATATGGAATGGAGGAGTACTGGTGTCCAGTTCACAAAG 3539
QY 3481 ACAGATCTTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGT 3540
Db 3540 ACAGATCTTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGT 3599
QY 3541 CTTTCTGTCAGTTTCAGTGAATGAAATGGCAATTCCTCCAGTCACTACAGTCA 3600
Db 3600 CTTTCTGTCAGTTTCAGTGAATGAAATGGCAATTCCTCCAGTCACTACAGTCA 3659
QY 3601 CTGGTAGGCCCCAGTGAATGGCTAGCATTTAGCTGAATATGTTATTACAGAGAGAAC 3660
Db 3660 CTGGTAGGCCCCAGTGAATGGCTAGCATTTAGCTGAATATGTTATTACAGAGAGAAC 3719
QY 3661 AGCTTATCCTGAGTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3720
Db 3720 AGCTTATCCTGAGTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3779
QY 3721 GATAAATAGTTATTTTAAAGAACTAATTCCTGAACTGAACCTTAAATCATCAAGCAGAGTG 3780
Db 3780 GATAAATAGTTATTTTAAAGAACTAATTCCTGAACTGAACCTTAAATCATCAAGCAGAGTG 3839
QY 3781 GCCTCTACGTTTACTCCTTTGCTGAAAAAAA 3813
Db 3840 GCCTCTACGTTTACTCCTTTGCTGAAAAAAA 3872
```

RESULT 4
AAF63930
ID AAF63930 standard; DNA: 4406 BP.
XX

```
AC AAF63930;  
XX 05-APR-2001 (first entry)  
DE Human tankyrase2 clone consensus oligonucleotide SEQ ID NO: 106.  
XX Human: tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;  
KW inflammatory disorder; ds.  
XX Homo sapiens.  
XX WO200100849-A1.  
PN 04-JAN-2001.  
XX 28-JUN-2000; 2000WO-US17827.  
PF 29-JUN-1999; 99US-0141582.  
XX (ICOS-) ICOS CORP.  
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
PI WPI; 2001-102896/11.  
XX P-PSDB; AAB66290.  
XX New tankyrase2 polypeptides, useful for treating conditions mediated by  
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,  
PT inflammatory and autoimmune disorders -  
XX Example 2; Page 168-173; 242pp; English.  
XX The present invention provides the protein and coding sequence for the  
CC human tankyrase2 protein. This is found in two different versions,  
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
CC polyADP-ribosylation activity and is involved in the modification of  
CC TRF1, which is a telomere-specific binding protein. The regulation of  
CC telomere length, in which TRF1 has a role, is linked to ageing and  
CC cancer. The sequences are useful in the treatment of cancers and  
CC inflammatory disorders.  
XX Sequence 4406 BP; 1268 A; 910 C; 1054 G; 1174 T; 0 other;  
SQ  
Query Match 99.4%; Score 3791.4; DB 22; Length 4406;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
QY 1 CGCGCTGCTCCCGCCCGCGGCGAGCGGGGGGCGAGGAGCCAGGAGGGCGCGCG 60  
Db 67 CGCGCTGCTCCCGCCCGCGGCGAGCGGGGGGCGAGGAGCCAGGAGGGCGCGCG 126  
QY 61 TGGGCGCGCGCCATGGGACTGCGCGGATCGCGGTGACAGCAGGAGCCAAAGCGCCCGG 120  
Db 127 TGGGCGCGG-CCATGGGACTGCGCGGATCGCGGTGACAGCAGGAGCCAAAGCGCCCGG 185  
QY 121 CCTGAGCGGCTTCTCCGGGGGCGCTCGCCCTCTCTGCTCGCGGGGCGCGGGCTCTCTGC 180  
Db 186 CCTGAGCGGCTTCTCCGGGGGCGCTCGCCCTCTCTGCTCGCGGGGCGCGGGCTCTCTGC 245  
QY 181 TCCGGTGTGGCGCTGTTGCTGGCTGTGGCGCGCGCCAGGATCATGTGCGGTGCGCGCT 240  
Db 246 TCCGGTGTGGCGCTGTTGCTGGCTGTGGCGCGCGCCAGGATCATGTGCGGTGCGCGCT 305  
QY 241 GGGCGCGCGGGGAGCGGCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
Db 306 GCGCGCGCGGGGAGCGGCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 365  
QY 301 GAGAGCTGTTCGAGCGGTGCCGCAACGGGAGCTTGAAGAGTCAAGAGGCTGGTGACGC 360  
Db 366 GAGAGCTGTTCGAGCGGTGCCGCAACGGGAGCTTGAAGAGTCAAGAGGCTGGTGACGC 425  
QY 361 CTGAGAAGGTGAACAGCGCGCGAGGAGAAATCCACCCCGCTGCACCTTCGCGCG 420
```


Db 772 CCTGAGCGCTCTTCTCCGGGGGCTCGCCCTCTCTGCTCGGGGCGGGGCTCCTGC 831
QY 181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGCGCGCGCCAGGATCATGTGCGGCTCGCCGCT 240
Db 832 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGCGCGCGCCAGGATCATGTGCGGCTCGCCGCT 891
QY 241 GCCCCGCGGGGAGCGGCTGCGCGAGCGCGCGCGCGCGCGGCGGAGCGCGCGGCC 300
Db 892 GCCCCGCGGGGAGCGGCTGCGCGAGCGCGCGCGCGCGCGGCGGAGCGCGCGGCC 951
QY 301 GAGAGCTGTTTCGAGCGGCTGCGCGAACGGGAGCGTGGAAAGAGTCAAGAGCGTGTGACGC 360
Db 952 GAGAGCTGTTTCGAGCGGCTGCGCGAACGGGAGCGTGGAAAGAGTCAAGAGCGTGTGACGC 1011
QY 361 CTGAGAAGGTGAACGCCGCGACACCGCGCGGAGGAAATCCACCGCGCTGCATCTCGCG 420
Db 1012 CTGAGAAGGTGAACGCCGCGGACACCGCGGAGGAAATCCACCGCGCTGCATCTCGCGG 1071
QY 421 CAGGTTTTGGCGGAAAGACGTAGTTGAATATATTGCTTCAGAAATGGTGCAAATGTCCAAG 480
Db 1072 CAGGTTTTGGCGGAAAGACGTAGTTGAATATATTGCTTCAGAAATGGTGCAAATGTCCAAG 1131
QY 481 CAGCTGATGATGGGGCCTTATTCCTTTCATTAAGCATGCTCTTTTGGTCATGCTGAAG 540
Db 1132 CAGCTGATGATGGGGCCTTATTCCTTTCATTAAGCATGCTCTTTTGGTCATGCTGAAG 1191
QY 541 TAGTCAATCTCCTTTTGCACATGTTGCGACGTCGACGCCCAATGCTCGAGATAATTGGAATTATA 600
Db 1192 TAGTCAATCTCCTTTTGGCGCATGTTGCGACGCCCAATGCTCGAGATAATTGGAATTATA 1251
QY 601 CTCCTCTCCATGAAGTGCAAATTAAGGAAAGATTGATGTTTGCAATTGTGCTTTACAGC 660
Db 1252 CTCCTCTCCATGAAGTGCAAATTAAGGAAAGATTGATGTTTGCAATTGTGCTTTACAGC 1311
QY 661 ATGAGCTGAGCCAAACCATCCGAAATACAGATGGAAGGACACATTTGGAATTTAGCAGATC 720
Db 1312 ATGAGCTGAGCCAAACCATCCGAAATACAGATGGAAGGACACATTTGGAATTTAGCAGATC 1371
QY 721 CATCTGCCAAAGCAGTCTTACTGGTGAATATAAGAAAGATGAATCTTTAGAAAGTGCCA 780
Db 1372 CATCTGCCAAAGCAGTCTTACTGGTGAATATAAGAAAGATGAATCTTTAGAAAGTGCCA 1431
QY 781 GGAGTGGCAATGAAGAAAAATGATGGCTCTACTCACACCATTAATGTCACCTGCCAGC 840
Db 1432 GGAGTGGCAATGAAGAAAAATGATGGCTCTACTCACACCATTAATGTCACCTGCCAGC 1491
QY 841 CAAGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db 1492 CAAGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 1551
QY 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGCTCATGCTAAAGATAAAGCTGATCTGG 960
Db 1552 TTGTACAGCTGTTACTGCAACATGGAGCTGATGCTCATGCTAAAGATAAAGCTGATCTGG 1611
QY 961 TACCATTACACAATGCCTGTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db 1612 TACCATTACACAATGCCTGTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1671
QY 1021 ATGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTCAATGAGGAGCTT 1080
Db 1672 ATGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTCAATGAGGAGCTT 1731
QY 1081 CTAAAGAACAGGTTGAAGTATGTTCTCTCTTAAAGTTATGGTGCAGACCCCAACACTGC 1140
Db 1732 CTAAAGAACAGGTTGAAGTATGTTCTCTCTTAAAGTTATGGTGCAGACCCCAACACTGC 1791
QY 1141 TCAATTGTCAATAAAAAGTGTATAGACTTGGCTCCCAACACAGATTAAAAGAAAGAT 1200
Db 1792 TCAATTGTCAATAAAAAGTGTATAGACTTGGCTCCCAACACAGATTAAAAGAAAGAT 1851
QY 1201 TACCATATGAATTTAAAGGCCACTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACATC 1260
Db 1852 TAGCATATGAATTTAAAGGCCACTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTC 1911

QY 1261 GAATCAAAAAACATCTCTCTGGAATGGTGAATTTCAAGCATCCTCAAAACACATGAAA 1320
Db 1912 GAATCAAAAAACATCTCTCTGGAATGGTGAATTTCAAGCATCCTCAAAACACATGAAA 1971
QY 1321 CAGCATTCATTTGCTGCTGCATCTCCATATCCAAAAGAAAGCAATATGTGAACCTGT 1380
Db 1972 CAGCATTCATTTGCTGCTGCATCTCCATATCCAAAAGAAAGCAATATGTGAACCTGT 2031
QY 1381 TGCTAAGAAAAGGAGCAACATCAATGAAAAGACTAAAGAATTTCTGACTCCTCTGCACG 1440
Db 2032 TGCTAAGAAAAGGAGCAACATCAATGAAAAGACTAAAGAATTTCTGACTCCTCTGCACG 2091
QY 1441 TGGCATCTGAAAAGCTCATTAATGATGTTGTTGAAGTAGTGGTGAACATGAACAAGG 1500
Db 2092 TGGCATCTGAAAAGCTCATTAATGATGTTGTTGAAGTAGTGGTGAACATGAACAAGG 2151
QY 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATC 1560
Db 2152 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATC 2211
QY 1561 TACAACCTCGCGCTACTCCTGAGCTATGGTGTGATCCCTAACATTTATATCCCTTCAGG 1620
Db 2212 TACAACCTCGCGCTACTCCTGAGCTATGGTGTGATCCCTAACATTTATATCCCTTCAGG 2271
QY 1621 GCTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCCTCCAAGAGGSTATCT 1680
Db 2272 GCTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCCTCCAAGAGGSTATCT 2331
QY 1681 CATTAGSTAAATTCAGAGGACAGACACAATTTGCTGGAGCTGCAAAAGCTGGAGATGTCG 1740
Db 2332 CATTAGSTAAATTCAGAGGACAGACACAATTTGCTGGAGCTGCAAAAGCTGGAGATGTCG 2391
QY 1741 AAACGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTC 1800
Db 2392 AAACGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTC 2451
QY 1801 AGTCTACACCACTTTCATTTTGCAGCTGGTATACAGAGTGTCCGTGGTGAATATCTGC 1860
Db 2452 AGTCTACACCACTTTCATTTTGCAGCTGGTATACAGAGTGTCCGTGGTGAATATCTGC 2511
QY 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGCGCTTGTACCTTTGCACAATG 1920
Db 2512 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGCGCTTGTACCTTTGCACAATG 2571
QY 1921 CATGTTCTTATGGACATTTATGAAGTTGCAGAACCTTCTTGTTTAAACATGGAGCAGTAGTTA 1980
Db 2572 CATGTTCTTATGGACATTTATGAAGTTGCAGAACCTTCTTGTTTAAACATGGAGCAGTAGTTA 2631
QY 1981 ATGTAGCTGATTTTATGGAAATTTACACCTTTTACATGAAGCAGCAGCAAAAGAAAATATG 2040
Db 2632 ATGTAGCTGATTTTATGGAAATTTACACCTTTTACATGAAGCAGCAGCAAAAGAAAATATG 2691
QY 2041 AAAATTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTTACCACAAAACAGGATGSA 2100
Db 2692 AAAATTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTTACACAAAACAGGATGSA 2751
QY 2101 ATACTCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAG 2160
Db 2752 ATACTCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAG 2811
QY 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTATAGCCAGAGTGAAGAAGTTGTCTT 2220
Db 2812 ATGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTATAGCCAGAGTGAAGAAGTTGTCTT 2871
QY 2221 CTCCTGATAATGTAAATTTGCCGCGATACCCAAAGGACAGACATTTCAACACCTTTACATTTAG 2280
Db 2872 CTCCTGATAATGTAAATTTGCCGCGATACCCAAAGGACAGACATTTCAACACCTTTACATTTAG 2931
QY 2281 CAGCTGGTTATAATTAATTTAGAAAGTTGCAGAGTATTTGTTTACAAACAGGAGCTGATGTGA 2340
Db 2932 CAGCTGGTTATAATTAATTTAGAAAGTTGCAGAGTATTTGTTTACAAACAGGAGCTGATGTGA 2991

QY 2341 ATCCCAAGCAAAAGGAGGACTTATTCCTTTACATTAATGACGATCTTACGGGCATGTAG 2400
Db ATCCCAAGCAAAAGGAGGACTTATTCCTTTACATTAATGACGATCTTACGGGCATGTAG 3051
QY 2401 ATGTACAGCTCTACTTAATTAAGTATAATGCAATGTCATGTCACGACGACAAATGGCCTT 2460
Db ATGTACAGCTCTACTTAATTAAGTATAATGCAATGTCATGTCACGACGACAAATGGCCTT 3111
QY 2461 TCACACCTTTGACGAAGCAGCCCAAAAGGAGCAACACAGCTTTTGTGCTTGTGCTAG 2520
Db TCACACCTTTGACGAAGCAGCCCAAAAGGAGCAACACAGCTTTTGTGCTTGTGCTAG 3171
QY 2521 CCATGGAGCTGACCCGACTCTTTAAATAATCAGGAAGCAACACACTTTAGATTAGTTT 2580
Db CCATGGAGCTGACCCGACTCTTTAAATAATCAGGAAGCAACACACTTTAGATTAGTTT 3231
QY 2581 CAGCGATGATGTCAGGGCTCTCTGACAGCAGCCATGCCCCATCTGCTCTGCCCTCTT 2640
Db CAGCGATGATGTCAGGGCTCTCTGACAGCAGCCATGCCCCATCTGCTCTGCCCTCTT 3291
QY 2641 GTTACAAGCTCAAGTCTCAATGGTGTGAGAAAGCCAGGAGCACTGCAGATGCTCTCT 2700
Db GTTACAAGCTCAAGTCTCAATGGTGTGAGAAAGCCAGGAGCACTGCAGATGCTCTCT 3351
QY 2701 CTTACAGCTCCTAGCCCATCAAGCTTTCTGACGCCAGCAGCTTTGACAACTTATCTG 2760
Db CTTACAGCTCCTAGCCCATCAAGCTTTCTGACGCCAGCAGCTTTGACAACTTATCTG 3411
QY 2761 GGAGTTTTCAGAACTCTCTCAGTAGTGTAGTTCAAGTGAAGGAGGCTTCCAGTT 2820
Db GGAGTTTTCAGAACTCTCTCAGTAGTGTAGTTCAAGTGAAGGAGGCTTCCAGTT 3471
QY 2821 TGGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATTAAGTCAATTCGTAAGGAATCTTG 2880
Db TGGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATTAAGTCAATTCGTAAGGAATCTTG 3531
QY 2881 GACTTGAGCACCCTAATGGATATATTTAGAGAGAAACAGATCACTTTGGATGTAATAGTTG 2940
Db GACTTGAGCACCCTAATGGATATATTTAGAGAGAAACAGATCACTTTGGATGTAATAGTTG 3591
QY 2941 AGATGGGACAAAGAGCTGAAGGAGATGGAATCAATCTGATGACATGACATGACCAAAAC 3000
Db AGATGGGACAAAGAGCTGAAGGAGATGGAATCAATCTGATGACATGACATGACCAAAAC 3651
QY 3001 TAATTAAGGAGTCGAGAGACTTATCTCCGGCAACAAGCTCTTAACCCATATTTAACTT 3060
Db TAATTAAGGAGTCGAGAGACTTATCTCCGGCAACAAGCTCTTAACCCATATTTAACTT 3711
QY 3061 TGAACACCTCTGTGAGTGAACAATTTCTATAGATCTGCTCTGATGATAAAGAGTTTC 3120
Db TGAACACCTCTGTGAGTGAACAATTTCTATAGATCTGCTCTGATGATAAAGAGTTTC 3771
QY 3121 AGTCTGTGGAGGAGAGATCCAAAGTACAGTTCGAGAGCAGAGATGAGGTCATGCGAG 3180
Db AGTCTGTGGAGGAGAGATCCAAAGTACAGTTCGAGAGCAGAGATGAGGTCATGCGAG 3772
QY 3181 GTGGAATCTTCAACAGATACATATTTCTCAAGATTGAGAGGTTTCTAACAAGAACTAT 3240
Db GTGGAATCTTCAACAGATACATATTTCTCAAGATTGAGAGGTTTCTAACAAGAACTAT 3891
QY 3241 GGGAAAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAACCAACCAACCATGCCAATG 3300
Db GGGAAAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAACCAACCAACCATGCCAATG 3951
QY 3301 AACGAATGCTATTTTATGGGTCTCTTTTGTGAATGCAATATTCACAAAGGTTTGATG 3360
Db AACGAATGCTATTTTATGGGTCTCTTTTGTGAATGCAATATTCACAAAGGTTTGATG 4011
QY 3361 AAAGGATCGGTACATAGTGTGATTTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTT 3420
Db AAAGGATCGGTACATAGTGTGATTTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTT 4071
QY 3421 CCAAAAGCAATCAATATGATATGGAATTTGGAGGAGGTACTGGGTGTCAGATTCACAAAG 3480

Db 4072 CCAAAAGCAATCAATATGATATGGAATTTGGAGGAGGTACTGGGTGTCAGATTCACAAAG 4131
QY 3481 ACAGATCTTTGTACATTTGCCACAGCAGCTGCTCTTTTCCGGGTAAACCTTGGAAAGT 3540
Db 4132 ACAGATCTTTGTACATTTGCCACAGCAGCTGCTCTTTTCCGGGTAAACCTTGGAAAGT 4191
QY 3541 CTTTCTGTCAGTTTCAAGTGAATGAAATGGCAATTTCTCTCCAGGTCATCACTCAGTCA 3600
Db 4192 CTTTCTGTCAGTTTCAAGTGAATGAAATGGCAATTTCTCTCCAGGTCATCACTCAGTCA 4251
QY 3601 CTGGTAGGCCAGCTGTAATGCGCTAGCATAGCTGAATGTAATTTACAGAGGAGAAC 3660
Db 4252 CTGGTAGGCCAGCTGTAATGCGCTAGCATAGCTGAATGTAATTTACAGAGGAGAAC 4311
QY 3661 AGCTTATCTGATGATTTAAATTTACATTTACAGATTTAGAGGCTGAGGTATGTCGATG 3720
Db 4312 AGCTTATCTGATGATTTAAATTTACATTTACAGATTTAGAGGCTGAGGTATGTCGATG 4371
QY 3721 GATAAATAGTTTATTTAAGAACTAATTTCCACTGAACCTAAAATCATCAAGCAGCAGTG 3780
Db 4372 GATAAATAGTTTATTTAAGAACTAATTTCCACTGAACCTAAAATCATCAAGCAGCAGTG 4431
QY 3781 GCCTCTACGTTTACTCTCTTGTGTAAGAAAAA 3813
Db 4432 GCCTCTACGTTTACTCTCTTGTGTAAGAAAAA 4464

RESULT 6
AAF63952
ID AAF63952 standard; DNA; 5002 BP.
XX
AC AAF63952;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 TANK2-LONG coding sequence SEQ ID NO: 132.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17827.
XX
PR 29-JUN-1999; 99US-0141582.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
DR WPI; 2001-102896/11.
XX
DR P-PSDB; AAB66294.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX
PS Claim 5; Page 185-190; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
XX inflammatory disorders.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX
XX Example 1; Page 158-162; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
SQ Sequence 4127 BP; 1245 A; 813 C; 934 G; 1135 T; 0 other:
Query Match 94.2%; Score 3595.8; DB 22; Length 4127;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3597; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 215 GGCAGGATCATGTGGGTGCGCGTGGCGGGGGAGCGGCGTGGCGAGCGCGC 274
Db 1 GGCAGGATCATGTGGGTGCGCGTGGCGGGGGAGCGGCGTGGCGAGCGCGC 60
QY 275 GGCAGGCGGTGGAGCGCGCGCGGAGAGCTGTTGAGGCGTGGCGCAACGGGAGCT 334
Db 61 GGCAGGCGGTGGAGCGCGCGCGGAGAGCTGTTGAGGCGTGGCGCAACGGGAGCT 120
QY 335 GGAAGAGTCAAGAGCGGTGACGCCGTGAGAAGTGAACCGCGCACCGCGGGCGAG 394
Db 121 GGAAGAGTCAAGAGCGGTGAGCGCTGAGAGGTGAACCGCGCACCGCGGGCGAG 180
QY 395 GAAATCCACCGCGTGCACCTTGGCGGAGGTTTGGGCGGAAAGACGAGTGAATATT 454
Db 181 GAAATCCACCGCGTGCACCTTGGCGGAGGTTTGGGCGGAAAGACGAGTGAATATT 240
QY 455 GCTTCAGAGTGGCAATGTCAGAGCGAGTATGATGGGCGCTTATTCCTCTTCATAA 514
Db 241 GCTTCAGAGTGGCAATGTCAGAGCGAGTATGATGGGCGCTTATTCCTCTTCATAA 300
QY 515 TGCATGCTCTTTTGGTGCATGCTGAAGTGTCAATCTCTTTTGGGACATGGTCAGACCC 574
Db 301 TGCATGCTCTTTTGGTGCATGCTGAAGTGTCAATCTCTTTTGGGACATGGTCAGACCC 360
QY 575 CAATGCTCGAGATTAATGGAATTATACCTCCATCAAGCTGCAATTAAAGAAAGAT 634
Db 361 CAATGCTCGAGATTAATGGAATTATACCTCCATCAAGCTGCAATTAAAGAAAGAT 420
QY 635 TGATGTTTGCATGCTGTTACAGCATGGAGCTGAGCAACATCCGAAATACAGATGG 694
Db 421 TGATGTTTGCATGCTGTTACAGCATGGAGCTGAGCAACATCCGAAATACAGATGG 480
QY 695 AAGGACAGCATTTGATTTAGCAGATCCATCTGCCAAGCAGTGCCTTACTGGTGAATATA 754
Db 481 AAGGACAGCATTTGATTTAGCAGATCCATCTGCCAAGCAGTGCCTTACTGGTGAATATA 540
QY 755 GAAAGTGAACCTTTAGAAAGTGCAGAGTGGCAATGAAGAAATAATGATGGCTCTACT 814
Db 541 GAAAGTGAACCTTTAGAAAGTGCAGAGTGGCAATGAAGAAATAATGATGGCTCTACT 600
QY 815 CACACCATTAATGTCAACTGCCAGCAGTGTGATGGCAAGTCAACTCCATTAACATTT 874
Db 601 CACACCATTAATGTCAACTGCCAGCAGTGTGATGGCAAGTCAACTCCATTAACATTT 660
QY 875 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTTACTGCAACATGGAGCTGATGT 934
Db 661 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTTACTGCAACATGGAGCTGATGT 720
QY 935 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACATGCTCTTCTTATGGTCAATTA 994
Db 721 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACATGCTCTTCTTATGGTCAATTA 780

QY 995 TGAAGTAACTGAACCTTTTGGTCAAGCATGGTGCCTGTAAATGCAATGCACTTGTGGCA 1054
Db 781 TGAAGTAACTGAACCTTTTGGTCAAGCATGGTGCCTGTAAATGCAATGCACTTGTGGCA 840
QY 1055 ATTCACTCTCTTTTCATGAGCGAGCTTCTAAGAACAGGGTGAAGTATGTTCTTCTCTT 1114
Db 841 ATTCACTCTCTTTTCATGAGCGAGCTTCTAAGAACAGGGTGAAGTATGTTCTTCTCTT 900
QY 1115 AGTTTATGGTGCAGAGCCCAACACTGCTCAATTTGTCAATAAAGTGTATAGACTTGGC 1174
Db 901 AGTTTATGGTGCAGAGCCCAACACTGCTCAATTTGTCAATAAAGTGTATAGACTTGGC 960
QY 1175 TCCACACACACAGTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTGTGCTGA 1234
Db 961 TCCACACACACAGTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTGTGCTGA 1020
QY 1235 AGTGCACGAGAAGCTGATGTTTACTCGAATCAAAAAACATCTCTCTCGAATGTGAA 1294
Db 1021 AGTGCACGAGAAGCTGATGTTTACTCGAATCAAAAAACATCTCTCTCGAATGTGAA 1080
QY 1295 TTTCAAGCATCTCAACACATGAAACAGCATTTGCATTTGTGCTGCTGCCATATCC 1354
Db 1081 TTTCAAGCATCTCAACACATGAAACAGCATTTGCATTTGTGCTGCTGCCATATCC 1140
QY 1355 CAAAAAGAACCAATATGTGAAGTGTGTGAAGAAAGAGCAACATCAATGAAAAGAC 1414
Db 1141 CAAAAAGAACCAATATGTGAAGTGTGTGAAGAAAGAGCAACATCAATGAAAAGAC 1200
QY 1415 TAAAGAAATTTTGAAGTGTGTGCTGCTGCAAGTGTGCAATTAATGATGTTGTTGA 1474
Db 1201 TAAAGAAATTTTGAAGTGTGTGCTGCTGCAAGTGTGCAATTAATGATGTTGTTGA 1260
QY 1475 AGTAGTGTCAACATGAAGCAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1534
Db 1261 AGTAGTGTCAACATGAAGCAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1320
QY 1535 ACAGAGAGCTGCAATATTTGCTGCTATCAAAAGCTGCGGCTACTCTCTGAGCTATGGGTG 1594
Db 1321 ACAGAGAGCTGCAATATTTGCTGCTATCAAAAGCTGCGGCTACTCTCTGAGCTATGGGTG 1380
QY 1595 TGATCTTAACATTTATATCCCTTCAGGCTTTACTGCTTTACAGATGGGAAATGATGCT 1654
Db 1381 TGATCTTAACATTTATATCCCTTCAGGCTTTACTGCTTTACAGATGGGAAATGATGCT 1440
QY 1655 ACAGCAACTCTCTCAAGAGGTTATCTCATTAGGTAAATTCAGAGCGACAGACAATGCT 1714
Db 1441 ACAGCAACTCTCTCAAGAGGTTATCTCATTAGGTAAATTCAGAGCGACAGACAATGCT 1500
QY 1715 GGAAGCTGCAAGAGCTGGAGATGTCGAAACTGTAAAAAACTGTGTACTGTTGAGAGTGT 1774
Db 1501 GGAAGCTGCAAGAGCTGGAGATGTCGAAACTGTAAAAAACTGTGTACTGTTGAGAGTGT 1560
QY 1775 CAATGTCAGAGACATTTGAAGGGCTGAGTACACCACTTCATTTTCAGCTGGGTATAA 1834
Db 1561 CAATGTCAGAGACATTTGAAGGGCTGAGTACACCACTTCATTTTCAGCTGGGTATAA 1620
QY 1835 CAGAGTGTCCGTGGTGAATATCTGCTACAGATGGAGCTGATGCTGATGCTAAAGATAA 1894
Db 1621 CAGAGTGTCCGTGGTGAATATCTGCTACAGATGGAGCTGATGCTGATGCTAAAGATAA 1680
QY 1895 AGGAGGCTTGTACCTTTGCAACATGATGTTTATGGACATTTAAGTGTGAGAACT 1954
Db 1681 AGGAGGCTTGTACCTTTGCAACATGATGTTTATGGACATTTAAGTGTGAGAACT 1740
QY 1955 TCTTTGTTAAACATGGAGCAGTAGTAAATGATGCTGATTTATGGAATTTACACCTTTACA 2014
Db 1741 TCTTTGTTAAACATGGAGCAGTAGTAAATGATGCTGATTTATGGAATTTACACCTTTACA 1800
QY 2015 TGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGAGA 2074
Db 1801 TGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGAGA 1860

```
QY 2075 CCCTACCAAAAAACAGGGATGAAATACCTCTTTGGATCTTGTTAAAGATGGAGATAC 2134
Db 1861 CCCACAAAAAACAAGGATGAAATACCTCTTTGGATCTTGTTAAAGATGGAGATAC 1920
QY 2135 AGATATTCAAGATCTGCTTAGGGAGATGACAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 2194
Db 1921 AGATATTCAAGATCTGCTTAGGGAGATGACAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 1980
QY 2195 TTTAGCCAGAGCTGAAGAAGTTGCTTCTCCCTGATAAATCTAAATTCGCCGCGATACCCRAAG 2254
Db 1981 TTTAGCCAGAGCTGAAGAAGTTGCTTCTCCCTGATAAATCTAAATTCGCCGCGATACCCRAAG 2040
QY 2255 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAAGTTGCAGAGTA 2314
Db 2041 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAAGTTGCAGAGTA 2100
QY 2315 TTTCTTACACACAGGAGCTGATGTGAATGTCGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2374
Db 2101 TTTCTTACACACAGGAGCTGATGTGAATGTCGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2160
QY 2375 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTATAAAGTATAATGTCATG 2434
Db 2161 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTATAAAGTATAATGTCATG 2220
QY 2435 TGTCAAATGCCACGACAAATGGGCTTTACACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2494
Db 2221 TGTCAAATGCCACGACAAATGGGCTTTACACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2280
QY 2495 AACACAGCTTTGCTGCTTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2554
Db 2281 AACACAGCTTTGCTGCTTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2340
QY 2555 AGGACAAACACCTTTAGATTAGTTTTCAGCGGATGTGTGAGCGCTCTTCTGCAGCAGC 2614
Db 2341 AGGACAAACACCTTTAGATTAGTTTTCAGCGAGATGTGTGAGCGCTCTTCTGCAGCAGC 2400
QY 2615 CATGCCCCCATCTGCTCTGCCCTCTGTTTACAGCCCTCAAGTGTCAATGGTGTGAAG 2674
Db 2401 CATGCCCCCATCTGCTCTGCCCTCTGTTTACAGCCCTCAAGTGTCAATGGTGTGAAG 2460
QY 2675 CCCAGGAGCCACTGCAGATGCTCTCTCTCAGGTGCCATCTAGCCCATCAAGCCCTTCTGC 2734
Db 2461 CCCAGGAGCCACTGCAGATGCTCTCTCTCAGGTGCCATCTAGCCCATCAAGCCCTTCTGC 2520
QY 2735 AGCCAGCAGTCTTGACAACTTATCTGGGAGCTTTTTCAGAACTGTCTTTCAGTAGTTTC 2794
Db 2521 AGCCAGCAGTCTTGACAACTTATCTGGGAGCTTTTTCAGAACTGTCTTTCAGTAGTTTC 2580
QY 2795 AAGTGGAAACAGAGGCTGCTTCCAGTTTGGAGAAAGAGGAGGTTCCAGGAGTAGATTAG 2854
Db 2581 AAGTGGAAACAGAGGCTGCTTCCAGTTTGGAGAAAGAGGAGGTTCCAGGAGTAGATTAG 2640
QY 2855 CATAACTCAATTGCTAAGGAATCTTGGACTTGGACCTTAATGGATATATTTCAGAGAGA 2914
Db 2641 CATAACTCAATTGCTAAGGAATCTTGGACTTGGACCTTAATGGATATATTTCAGAGAGA 2700
QY 2915 ACAGATCACTTTGGATGTATTAGTTGAGATGGGCGACAAGGAGCTGAAGGAGATTGGAAT 2974
Db 2701 ACAGATCACTTTGGATGTATTAGTTGAGATGGGCGACAAGGAGCTGAAGGAGATTGGAAT 2760
QY 2975 CAATGCTTATGGACATAGGCACAACTAATTAAGGAGCTCGAGAGCTTATCTCCGACA 3034
Db 2761 CAATGCTTATGGACATAGGCACAACTAATTAAGGAGCTCGAGAGCTTATCTCCGACA 2820
QY 3035 ACAAGTCTTAACCCATATTAACTTTGAACACCTCTGTTAGTGGAACTTCTTATAGA 3094
Db 2821 ACAAGTCTTAACCCATATTAACTTTGAACACCTCTGTTAGTGGAACTTCTTATAGA 2880
QY 3095 TCTGCTCTGATGATAAAGATTTTCAGTCTGTGGAGGAAGAGATGCAAGTACAGTTCC 3154
Db 2881 TCTGCTCTGATGATAAAGATTTTCAGTCTGTGGAGGAAGAGATGCAAGTACAGTTCC 2940
QY 3155 AGAGCACAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATCTCAAGAT 3214
```

```
Db 2941 AGAGCACAGATGGAGTCTATGCAGGTGGATCTTCAACAGATACAATATCTCAAGAT 3000
QY 3215 TCAGAAGGTTTGTAAACAAGAACTATGGAAAGATPACACTCACCGGAGAAAGAGTTTC 3274
Db 3001 TCAGAAGGTTTGTAAACAAGAACTATGGAAAGATPACACTCACCGGAGAAAGAGTTTC 3060
QY 3275 TGAAGAAACACAAACCATGCCAATGAACGAATGCTATTTCATGGGTCCTCTTTGTGAA 3334
Db 3061 TGAAGAAACACAAACCATGCCAATGAACGAATGCTATTTCATGGGTCCTCTTTGTGAA 3120
QY 3335 TGCAATTATCCACAAAGGCTTTGATGAAAGCATGCGTACATAGTGTATGTTTGGAGC 3394
Db 3121 TGCAATTATCCACAAAGGCTTTGATGAAAGCATGCGTACATAGTGTATGTTTGGAGC 3180
QY 3395 TGGCATTATTTTCTGCTGAAACTCTTCCAAAGCAATCAATATATATGCAATTTGGAGG 3454
Db 3181 TGGCATTATTTTCTGCTGAAACTCTTCCAAAGCAATCAATATATATGCAATTTGGAGG 3240
QY 3455 AGGTACTGGGTGTCAGTTCCAAAGACAGATCTTGTACATTTGCCACAGCAGCTGCT 3514
Db 3241 AGGTACTGGGTGTCAGTTCCAAAGACAGATCTTGTACATTTGCCACAGCAGCTGCT 3300
QY 3515 CTTTTCGCCGGTAACCTTTGGGAAAGTCTTCTCCTGCAGTTCAAGTCAATGAAATGGCACA 3574
Db 3301 CTTTTCGCCGGTAACCTTTGGGAAAGTCTTCTCCTGCAGTTCAAGTCAATGAAATGGCACA 3360
QY 3575 TTCTCCTCCAGGTCATCAGTCAGTCACTGGTAGGCCAGTGTAAATGGCCTTAGCATTAGC 3634
Db 3361 TTCTCCTCCAGGTCATCAGTCAGTCACTGGTAGGCCAGTGTAAATGGCCTTAGCATTAGC 3420
QY 3635 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGCAGTTCAAGTCAATGAAATGGCACA 3694
Db 3421 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGCAGTTCAAGTCAATGAAATGGCACA 3480
QY 3695 TATGAGCCCTCAAGGATGTTGATGGATAAAATAGTTATTTTAAAGAACTAATTTCCACTG 3754
Db 3481 TATGAGCCCTCAAGGATGTTGATGGATAAAATAGTTATTTTAAAGAACTAATTTCCACTG 3540
QY 3755 AACCTAAATCATCAAGACAGCAGTGGCCCTAGCTTTTACTCCTTTGCTGAAAAAAA 3813
Db 3541 AACCTAAATCATCAAGACAGCAGTGGCCCTAGCTTTTACTCCTTTGCTGAAAAAAA 3599

RESULT 8
AAD02578
ID AAD02578 standard; DNA; 4512 Bp.
XX
AC AAD02578;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human tankyrase homolog protein (THP) DNA.
XX
KW Human; tankyrase homolog protein; THP; gene therapy; cancer;
KW tumour; basal cell carcinoma; therapy; genetic mapping;
KW cytostatic; ds.
XX
OS Homo sapiens.
XX
FH Key
CDS 23..3523
FT /*tag= a
FT /product= "Human tankyrase homolog protein (THP)"
FT /note= "This region is specifically claimed in
XX claim 4 as SEQ ID No:4"
PN WO200104326-A1.
XX
PD 18-JAN-2001.
XX
PF 03-JUL-2000; 2000WO-BP06609.
XX
```

PR 09-JUL-1999; 99US-0350982.
XX (PHAA) PHARMACIA & UPOJOHN SPA.
XX
XX Berthelsen J, Toma S, Isacchi A;
PI
XX WPI; 2001-168422/17.
DR P-PSDB; AAY72589.
DR
XX
XX New tankyrase homolog protein (THP) polynucleotide and polypeptide
PT useful in gene therapy, diagnosis and treatment or prevention of
PT unregulated cell growth, such as cancer or tumor cell growth -
XX
XX Claim 1; Page 41-42; 60pp; English.
XX
XX The present sequence is a DNA coding for human tankyrase homolog
CC protein (THP). The THP polypeptides and polynucleotides of the
CC invention are useful in gene therapy and for treating or preventing
CC unregulated cell growth such as cancer or tumour (e.g. basal cell
CC carcinoma). The nucleic acid molecules of the invention and their
CC fragments are useful for restriction fragment length polymorphism
CC (RFLP) associated with certain disorders, as well as for genetic
CC mapping. Antisense oligonucleotides, or fragments of nucleic acid
CC encoding THP are useful as diagnostic tools for probing the
CC expression of THP gene in various tissues. THP can be used as
CC antigens for raising antibodies against them and in assays for
CC identifying compounds that modulate their activity. They are used in
CC the manufacture of a medicament directed towards cancers or tumours.
CC THP are also useful for screening compounds in a variety of drugs
CC screening techniques and as a research tool for identification,
CC characterisation and purification of interacting, regulatory proteins.
XX
XX Sequence 4512 BP; 1356 A; 878 C; 1002 G; 1268 T; 8 other;
SQ

Query Match 94.2%; Score 3594.8; DB 22; Length 4512;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3593; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 210 GCGGGCGGCGAGCATCATGTCGGGTCGCGCGCGGGGGAGCGGCTGCGCGAGC 269
Db
Qy 270 GCGGGCGGCGAGCGGCGGCGCGCGCGCGGAGCGTGTTCGAGGCGTGGCGCAACGGG 329
Db 68 GCGGGCGGCGAGCGGCGGCGCGCGCGCGGAGCGTGTTCGAGGCGTGGCGCAACGGG 127
Qy 330 GAGGTGGAACGAGTCAAGAGGCTGTGAGCGCTGAGAGGTTGACAGCGCGGACACGCG 389
Db 128 GAGGTGGAACGAGTCAAGAGGCTGTGAGCGCTGAGAGGTTGACAGCGCGGACACGCG 187
Qy 390 GCGAGGAAATCCACCGCGTGCACCTTCGCGCGAGCTTTTGGCGGAAAGAGCTAGTTCAA 449
Db 188 GCGAGGAAATCCACCGCGTGCACCTTCGCGCGAGCTTTTGGCGGAAAGAGCTAGTTCAA 247
Qy 450 TATTTCCTTCAGAAATGGTSCAAATGCCAAGCACGTCATGATGGGGGCTTATTCCTCTT 509
Db 248 TATTTCCTTCAGAAATGGTSCAAATGCCAAGCACGTCATGATGGGGGCTTATTCCTCTT 307
Qy 510 CATATGATGCTCTTTGGTTCATGCTGAGTGAAGTATGCTCTCTTTTGGGACATGGTGA 569
Db 308 CATATGATGCTCTTTGGTTCATGCTGAGTGAAGTATGCTCTCTTTTGGGACATGGTGA 367
Qy 570 GACCCCAATGCTCGAGATTAATGGAAATTAATCTCTCTCATGAAGTGCATTAAGGA 629
Db 368 GACCCCAATGCTCGAGATTAATGGAAATTAATCTCTCTCATGAAGTGCATTAAGGA 427
Qy 630 AAGATTGATGTTTGGCATTTGCTGTTTACAGCATGGAGCTGAGCCCAACCATCCGAATACA 689
Db 428 AAGATTGATGTTTGGCATTTGCTGTTTACAGCATGGAGCTGAGCCCAACCATCCGAATACA 487
Qy 690 GATGGAAGACACGATTTGATTTAGCAGATCCATCTGCCAAGACAGTCTTACTGGTGA 749
Db 488 GATGGAAGACACGATTTGATTTAGCAGATCCATCTGCCAAGACAGTCTTACTGGTGA 547

Qy 750 TATAAGAAAGATCAACTCTTAGAAAAGTCCAGGAGTGGCAATCAAGAAAAAATGATGGCT 809
Db 548 TATAAGAAAGATCAACTCTTAGAAAAGTCCAGGAGTGGCAATCAAGAAAAAATGATGGCT 607
Qy 810 CTACTCACACCATTAATGTCAACTGCCACGCAAGTATGGCAGAAAGTCAACTCCATTA 869
Db 608 CTACTCACACCATTAATGTCAACTGCCACGCAAGTATGGCAGAAAGTCAACTCCATTA 667
Qy 870 CATTTGGCAGCAGGATATACAGAGTAAAGATGTACAGCTGTTACTGCAACATGGAGCT 929
Db 668 CATTTGGCAGCAGGATATACAGAGTAAAGATGTACAGCTGTTACTGCAACATGGAGCT 727
Qy 930 GATGTCATGCTAAAGATAAAGTGTATGTTACCAATACACAATGCCTGTTCTTATGCT 989
Db 728 GATGTCATGCTAAAGATAAAGTGTATGTTACCAATACACAATGCCTGTTCTTATGCT 787
Qy 990 CATTTATCAAGTAATCAACTTTTGGTCAAGCATGGTGGCTGTGTAATCAATGCAATG 1049
Db 788 CATTTATCAAGTAATCAACTTTTGGTCAAGCATGGTGGCTGTGTAATCAATGCAATG 847
Qy 1050 TGGCAATTCACCTCTTCTCATGAGGCGAGCTTCTAAGAACAGGTTTGAAGTATGTTCTCT 1109
Db 848 TGGCAATTCACCTCTTCTCATGAGGCGAGCTTCTAAGAACAGGTTTGAAGTATGTTCTCT 907
Qy 1110 CTCTTAAGTTTATGTCAGACCCCAACACTGCTCAATTTGTCAATAAAAAGTGTATAGAC 1169
Db 908 CTCTTAAGTTTATGTCAGACCCCAACACTGCTCAATTTGTCAATAAAAAGTGTATAGAC 967
Qy 1170 TTGGCTCCACACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1229
Db 968 TTGGCTCCACACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1027
Qy 1230 CTGCAAGCTGCAGAGAGCTGATGTTACTCGAATCAAAAACATCTCTCTGGAATG 1289
Db 1028 CTGCAAGCTGCAGAGAGCTGATGTTACTCGAATCAAAAACATCTCTCTGGAATG 1087
Qy 1290 GTGAATTTCAAGCATCTCTCAACACATGAACAGCANTGATGCTGCTGCTATCTCCA 1349
Db 1088 GTGAATTTCAAGCATCTCTCAACACATGAACAGCANTGATGCTGCTGCTATCTCCA 1147
Qy 1350 TATCCCAAAAAGAAAGCAATATGTGAAGTGTGCTAGAAAAGGAGCAACATCAATGAA 1409
Db 1148 TATCCCAAAAAGAAAGCAATATGTGAAGTGTGCTAGAAAAGGAGCAACATCAATGAA 1207
Qy 1410 AAGACTAAAGAAATCTCTTGAAGTGTGCTAGAAAAGTGTGCTAGAAAAGTGTGCT 1469
Db 1208 AAGACTAAAGAAATCTCTTGAAGTGTGCTAGAAAAGTGTGCTAGAAAAGTGTGCT 1267
Qy 1470 GTTGAAGTGTGTTGAAGCAATGAAGCAAGGTTAATGCTCTGGATAATCTTGGTCAAGCT 1529
Db 1268 GTTGAAGTGTGTTGAAGCAATGAAGCAAGGTTAATGCTCTGGATAATCTTGGTCAAGCT 1327
Qy 1530 TCTCTACACAGAGCTGATATGTTGGTTCATCTACAAACCTGCCGCCCTACTCCTGAGCTAT 1589
Db 1328 TCTCTACACAGAGCTGATATGTTGGTTCATCTACAAACCTGCCGCCCTACTCCTGAGCTAT 1387
Qy 1590 GGGTGTGATCTCAACATTATATCCCTTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAA 1649
Db 1388 GGGTGTGATCTCAACATTATATCCCTTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAA 1447
Qy 1650 AATGTACAGCAACTCCTCCAAAGAGGTTATCTATTAGTAGTAATTCAGAGGAGACAGACAA 1709
Db 1448 AATGTACAGCAACTCCTCCAAAGAGGTTATCTATTAGTAGTAATTCAGAGGAGACAGACAA 1507
Qy 1710 TTGCTGGAGCTGCAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTTACAG 1769
Db 1508 TTGCTGGAGCTGCAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTTACAG 1567
Qy 1770 AGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACCTTCATTTTTCAGCTGGG 1829
Db 1568 AGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACCTTCATTTTTCAGCTGGG 1627

QY	1830	TATACAGAGTGTCCGTGGTGGAAATATCTGCTACAGCATGGAGCTGATGTGCATGCTATAA	1889
DB	1628	TATACAGAGTGTCCGTGGTGGAAATATCTGCTACAGCATGGAGCTGATGTGCATGCTATAA	1687
QY	1890	GATAAGGAGGCCCTTGCTACCTTTGCCAATGCATGTTCTTATGGACATATTACAATGTCGA	1949
DB	1688	GATTAAGPRGSCCTTGACCTTTGCACATGCATGTTCTTATGGACATATTGAAGTTGCA	1747
QY	1950	GAATCTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACAGCT	2009
DB	1748	GAATCTCTTGTTAAACATGCAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACAGCT	1807
QY	2010	TTACATGAAGCAGCAGCAAAAGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT	2069
DB	1808	TTACATGAAGCAGCAGCAAAAGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT	1867
QY	2070	GCAGACCTACAAAAAACAAGGATGAAATACTCTCTTTGGATCTGTTAAAGATGGA	2129
DB	1868	GCAGACCTACAAAAAACAAGGATGAAATACTCTCTTTGGATCTGTTAAAGATGGA	1927
QY	2130	GATACAGATATTCAGATCTGCTTAGGGAGATGCAGCTTTTGCCTAGATGCTGCCAAGAAG	2189
DB	1928	GATACAGATATTCAGATCTGCTTAGGGAGATGCAGCTTTTGCCTAGATGCTGCCAAGAAG	1987
QY	2190	GGTTGTTTAGCCAGGTGAAGAAGTTGCTCTCTCTGATTAATGTAATGTCGCGGATACC	2249
DB	1988	AGTTGTTTAGCCAGGTGAAGAAGTTGCTCTCTCTGATTAATGTAATGTCGCGGATACC	2047
QY	2250	CAAGGCAGACATTCACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA	2309
DB	2048	CAAGGCAGACATTCACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA	2107
QY	2310	GAGTATTGTTTACAACACGGAGCTGATGCAATGCCCAAGACAAGAGGAGCACTTATTCCT	2369
DB	2108	GAGTATTGTTTACAACACGGAGCTGATGCAATGCCCAAGACAAGAGGAGCACTTATTCCT	2167
QY	2370	TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT	2429
DB	2168	TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT	2227
QY	2430	GCATGTCTCAATGCCAGCACAAATGGGCTTTACACCTTTGCCAGGACGAGCCCAAG	2489
DB	2228	GCATGTCTCAATGCCAGCACAAATGGGCTTTACACCTTTGCCAGGACGAGCCCAAG	2387
QY	2490	GGAGAACACAGCTTTGTGCTTTGTGTAGCCCATGGAGCTGACCCGACTCTTTAAAAAT	2549
DB	2288	GGAGAACACAGCTTTGTGCTTTGTGTAGCCCATGGAGCTGACCCGACTCTTTAAAAAT	2347
QY	2550	CAGGAGGACAAACACCTTTAGATTTAGTTTTCAGCGGATGATGTCAGGCTCTCTGACA	2609
DB	2348	CAGGAGGACAAACACCTTTAGATTTAGTTTTCAGCGGATGATGTCAGGCTCTCTGACA	2407
QY	2610	GCAGCCATGCCCCATCTGCTCTGCCCTCTGTTTACAAGCCTCAAGTGCCTCAATGGTGTG	2669
DB	2408	GCAGCCATGCCCCATCTGCTCTGCCCTCTGTTTACAAGCCTCAAGTGCCTCAATGGTGTG	2467
QY	2670	AGRAGCCAGGAGCCACTGCCATGCTCTCTTTCAGTCCATCTAGCCCATCAAGCCTT	2729
DB	2468	AGRAGCCAGGAGCCACTGCCATGCTCTCTTTCAGTCCATCTAGCCCATCAAGCCTT	2527
QY	2730	CTGCGAGCCAGCAGTCTTGACAACTTATCTGGAGTTTTCAGAACTGCTCTTCAGTAGTT	2789
DB	2528	CTGCGAGCCAGCAGTCTTGACAACTTATCTGGAGTTTTCAGAACTGCTCTTCAGTAGTT	2587
QY	2790	AGTTCAAGTGGAAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGTTCCAGGAGTAGAT	2849
DB	2588	AGTTCAAGTGGAAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGTTCCAGGAGTAGAT	2647
QY	2850	TTTAGCATTAACCTCAATTCGTAAGGAATCTTGACTTGAGCACCTTAATGGATATATTGAG	2909
DB	2648	TTTAGCATTAACCTCAATTCGTAAGGAATCTTGACTTGAGCACCTTAATGGATATATTGAG	2707
QY	2910	AGAGAACAGATCACTTTGGATGATTAATGATGGGGACAAGGAGCTTGAAGGAGATT	2969

Db	2708	AGAGAACAGATCAC	TTTTGGATGATTAGTTGAGATGGGCACAGGAGCTC	AGGAGATT	2767
QY	2970	GGAAATCAATGCTTAT	GGACATAGCCACAAACTAA	TTAAAGAGTCGAGAGACTTAT	3029
Db	2768	GGAAATCAATGCTTAT	GGACATAGGCACAAACTAA	TTAAAGAGTCGAGAGACTTAT	2827
QY	3030	GGACAACAAGGTCCT	TAAACCATATTTAACTTTGAACACCTTC	TGGTACTGTGAACAAATTC	3089
Db	2828	GGACAACAAGGTCCT	TAAACCATATTTAACTTTGAACACCTTC	TGGTACTGTGAACAAATTC	2887
QY	3090	ATAGATCTGTCCTCT	GATGATAAAGAGTTTCAGTCTCT	TGGAGGAAGAGATGCAAACTACA	3149
Db	2888	ATAGATCTGTCCTCT	GATGATAAAGAGTTTCAGTCTCT	TGGAGGAAGAGATGCAAACTACA	2947
QY	3150	GTTTCGAGACACAGAGAT	CGAGGTCATGCGAGGTGGAA	CTTTCAACAGATACAATAT	3209
Db	2948	GTTTCGAGACACAGAGAT	CGAGGTCATGCGAGGTGGAA	CTTTCAACAGATACAATAT	3007
QY	3210	AAGATTGAGAAGGTTT	GTAAACAAGAACTATGGAAAGATACACT	CACCCGGAGAAAAGAA	3269
Db	3008	AAGATTGAGAAGGTTT	GTAAACAAGAACTATGGAAAGATACACT	CACCCGGAGAAAAGAA	3067
QY	3270	GTTTCTGAAGAAACAC	ACACCAATGCCAATGAACGAATGCT	TATTCATGGGTCTCCTTTT	3329
Db	3068	GTTTCTGAAGAAACAC	ACACCAATGCCAATGAACGAATGCT	TATTCATGGGTCTCCTTTT	3127
QY	3330	GTGAATGCAATTAT	CACAAAGGCTTTGATGAAAGGCATGCGT	ACATAGGTGATATGTTT	3389
Db	3128	GTGAATGCAATTAT	CACAAAGGCTTTGATGAAAGGCATGCGT	ACATAGGTGATATGTTT	3187
QY	3390	GGAGCTGGCATTTAT	TTTGTGCTGAAAACCTCTCCAAAGCAATCAAT	TATGTATGGAATT	3449
Db	3188	GGAGCTGGCATTTAT	TTTGTGCTGAAAACCTCTCCAAAGCAATCAAT	TATGTATGGAATT	3247
QY	3450	GGAGGAGGTACTGGG	TGTCACCAAGACAGATCTTGTTACATTTG	CCACAGGCAG	3509
Db	3248	GGAGGAGGTACTGGG	TGTCACCAAGACAGATCTTGTTACATTTG	CCACAGGCAG	3307
QY	3510	CTGCTCTTTTGGCGG	TACCTTGGGAAAGTCCTTCCAGCTTCAGTGC	CAATGAAATG	3569
Db	3308	CTGCTCTTTTGGCGG	TACCTTGGGAAAGTCCTTCCAGCTTCAGTGC	CAATGAAATG	3367
QY	3570	GCACATTTCCCTCC	AGGTCATCACCTAGTCATGTTAGCCCGAGTGAAT	GGCCCTAGCA	3629
Db	3368	GCACATTTCCCTCC	AGGTCATCACCTAGTCATGTTAGCCCGAGTGAAT	GGCCCTAGCA	3427
QY	3630	TTAGCTGAATATGTT	ATTATACAGAGGAGAACAGGCTTATCCCTGAGTAT	TAAATTAATCTTAC	3689
Db	3428	TTAGCTGAATATGTT	ATTATACAGAGGAGAACAGGCTTATCCCTGAGTAT	TAAATTAATCTTAC	3487
QY	3690	CAGATTATGAGCCCT	GGAAGGTATGGTTCGATGGATAAATAGTTATTTT	TAAAGAACTAATTC	3749
Db	3488	CAGATTATGAGCCCT	GGAAGGTATGGTTCGATGGATAAATAGTTATTTT	TAAAGAACTAATTC	3547
QY	3750	CACCTGAACCTTAA	AATCATCAAGCAGCAGTGGCCTCTACGTTTTT	TACTCCTTTGCTGAAAA	3809
Db	3548	CACCTGAACCTTAA	AATCATCAAGCAGCAGTGGCCTCTACGTTTTT	TACTCCTTTGCTGAAAA	3607
QY	3810	AAAA	3813		
Db	3608	AAAA	3611		

RESULT 9
AAAF63837
ID AAFAF63837 standard; DNA; 3508 bp.
XX
XX
XX AAFAF63837;
XX
XX
XX
DT 05-APR-2001 (first entry)
DE
DE Human tankyrase2 related coding sequence SEQ ID NO: 1

XX Human: tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX Homo sapiens.
OS W0200100849-A1.
PN W0200100849-A1.
XX 04-JAN-2001.
XX 28-JUN-2000; 2000WO-US17827.
XX 29-JUN-1999; 99US-0141582.
XX (ICOS-) ICOS CORP.
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
PI WPI: 2001-102896/11.
DR P-PSDB: AAB66278.
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX Disclosure; Page 105-109; 242pp; English.
XX The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
SQ Sequence 3508 bp; 1046 A; 710 C; 843 G; 909 T; 0 other;

Query Match 91.8%; Score 3504.8; DB 22; Length 3508;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3506; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 215 GCCCAGGATCATGTGGGTGCGCGTGGCGGGGAGCGGCGCTGCGGAGCGCGCGC 274
DB 1 GSCCAGGATCATGTGGGTGCGCGTGGCGGGGAGCGGCGCTGCGGAGCGCGCGC 60
QY 275 GCCCAGGCGGTGGAGCGCGCGCGCGAGAGCTGTTGAGCGCGTCCCGACACGGGAGCT 334
DB 61 GCCCAGGCGGTGGAGCGCGCGCGCGAGAGCTGTTGAGCGCGTCCCGACACGGGAGCT 120
QY 335 GGAACGAGTCAAGAGCGTGGTGACGCTGAGAAGGTGAACGCGCGACACGGCGGCGAG 394
DB 121 GGAACGAGTCAAGAGCGTGGTGACGCTGAGAAGGTGAACGCGCGACACGGCGGCGAG 180
QY 395 GAAATCCACCGCGTGCATTTGCGCGCGAGGTTTGGCGCGGAAGACGCTAGTTGAATATT 454
DB 181 GAAATCCACCGCGTGCATTTGCGCGCGAGGTTTGGCGCGGAAGACGCTAGTTGAATATT 240
QY 455 GCTTCAGATGTGCAATGTCCAGACGCTGATGATGGGCGCTTATTCCTCTTCATAA 514
DB 241 GCTTCAGATGTGCAATGTCCAGACGCTGATGATGGGCGCTTATTCCTCTTCATAA 300
QY 515 TGCATGCTCTTTTGGTTCATGCTGAAGTGTCAATCTCCTTTTGGCAGATGGTGCAGACCC 574
DB 301 TGCATGCTCTTTTGGTTCATGCTGAAGTGTCAATCTCCTTTTGGCAGATGGTGCAGACCC 360
QY 575 CAATGCTCAGATTAATTTGAATTAATTTACTCTCTCCATGAAGCTGCAATTAAGGAAAGAT 634
DB 361 CAATGCTCAGATTAATTTGAATTAATTTACTCTCTCCATGAAGCTGCAATTAAGGAAAGAT 420
QY 635 TGATGTTTCATTTGCTGCTTACAGCATGGAGCTGAGCCCAACCATCCGAAATACAGATGG 694
DB 421 TGATGTTTCATTTGCTGCTTACAGCATGGAGCTGAGCCCAACCATCCGAAATACAGATGG 480

QY 695 AAGGACGACATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGTCTTACTGGTGAATATA 754
DB 481 AAGGACGACATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGTCTTACTGGTGAATATA 540
QY 755 GAAAGATGAACCTTTAGAAAGTCCAGAGTGGCAGTGAATGAAGAAAAATGATGGCTCTACT 814
DB 541 GAAAGATGAACCTTTAGAAAGTCCAGAGTGGCAGTGAATGAAGAAAAATGATGGCTCTACT 600
QY 815 CACACCATTAATATGTCACCTGCCAGCAGTGCATGCGAGAAAGTCAACTCCATTTACATTT 874
DB 601 CACACCATTAATATGTCACCTGCCAGCAGTGCATGCGAGAAAGTCAACTCCATTTACATTT 660
QY 875 GGCAGCAGCATATAACAGAGTAAAGATTGTACAGCTGTGTACTGCAACATGGAGCTGATGT 934
DB 661 GGCAGCAGCATATAACAGAGTAAAGATTGTACAGCTGTGTACTGCAACATGGAGCTGATGT 720
QY 935 CCATGCTAAAGATAAAGTGTATCTGTACCACTTACAAATGCGCTGTCTTATGGTCAATTA 994
DB 721 CCATGCTAAAGATAAAGTGTATCTGTACCACTTACAAATGCGCTGTCTTATGGTCAATTA 780
QY 995 TGAAGTAACTGAACCTTTTGGTCAAGCATGTCCTGTGTAATGCAATGCACTTGTGGCA 1054
DB 781 TGAAGTAACTGAACCTTTTGGTCAAGCATGTCCTGTGTAATGCAATGCACTTGTGGCA 840
QY 1055 ATTCACTCTCTTTCATGAGGCGAGCTTCTAAGAACAGGGTTGAAGTATGTTCTTCTCTT 1114
DB 841 ATTCACTCTCTTTCATGAGGCGAGCTTCTAAGAACAGGGTTGAAGTATGTTCTTCTCTT 900
QY 1115 AAGTTATGGTGCAGAGCCCAACACTGCTCAATGTGCACATTAAGTGTCTATAGACTTGGC 1174
DB 901 AAGTTATGGTGCAGAGCCCAACACTGCTCAATGTGCACATTAAGTGTCTATAGACTTGGC 960
QY 1175 TCCCACACACAGCTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1234
DB 961 TCCCACACACAGCTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1020
QY 1235 AGTGCACGAGAGAGCTGATGTTACTCGAATCAAAAACATCTCTCTCTGGAATGTTGAA 1294
DB 1021 AGTGCACGAGAGAGCTGATGTTACTCGAATCAAAAACATCTCTCTCTGGAATGTTGAA 1080
QY 1295 TTTCAAGCATCTTCAACACATGAAACAGCATGTCATGTCGCTGCTGCACTCCCATATCC 1354
DB 1081 TTTCAAGCATCTTCAACACATGAAACAGCATGTCATGTCGCTGCTGCACTCCCATATCC 1140
QY 1355 CAAAAGAAACAAATATGTGAAGTGTGCTTAAGAAAGGAGCAACATCAATGAAAGAC 1414
DB 1141 CAAAAGAAACAAATATGTGAAGTGTGCTTAAGAAAGGAGCAACATCAATGAAAGAC 1200
QY 1415 TAAAGAAATTTGACTCTCTGACCTCTGACCTGCGCATCTGAGAAAGCTCAATGATGTTGA 1474
DB 1201 TAAAGAAATTTGACTCTCTGACCTCTGACCTGCGCATCTGAGAAAGCTCAATGATGTTGA 1260
QY 1475 AGTAGTGTGAAACATGAAGCAAGGTTAATGCTCTGGAATAATCTTGGTCAGACTTCTCT 1534
DB 1261 AGTAGTGTGAAACATGAAGCAAGGTTAATGCTCTGGAATAATCTTGGTCAGACTTCTCT 1320
QY 1535 ACACAGAGCTGCATATTTGTTGTTGTCATCTACAAACCTGCGCTACTCTCTGAGCTATGGTG 1594
DB 1321 ACACAGAGCTGCATATTTGTTGTTGTCATCTACAAACCTGCGCTACTCTCTGAGCTATGGTG 1380
QY 1595 TGATCTTAAATATATATCCCTTACGGCTTTTACTGCTTTACAGATGGGAATGAAATGT 1654
DB 1381 TGATCTTAAATATATATCCCTTACGGCTTTTACTGCTTTACAGATGGGAATGAAATGT 1440
QY 1655 ACAGCAACTCTCTCAAGAGGTTATCTCATTAAGTAAATTCAGAGGCGACAGACAATTTGCT 1714
DB 1441 ACAGCAACTCTCTCAAGAGGTTATCTCATTAAGTAAATTCAGAGGCGACAGACAATTTGCT 1500
QY 1715 GGAAGCTGCAAGGCTGGAGATGTGAAACCTGTAACAACTGTGTACTGTTGAGAGTGT 1774
DB 1501 GGAAGCTGCAAGGCTGGAGATGTGAAACCTGTAACAACTGTGTACTGTTGAGAGTGT 1560

QY 1775 CAACTGCAGAGACATTTGAAGGGCGTCAGCTACACCACCTTCATTTTCAGCTGGGTATAA 1834
Db 1561 CAACTGCAGAGACATTTGAAGGGCGTCAGCTACACCACCTTCATTTTCAGCTGGGTATAA 1620
QY 1835 CAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTCATGCTAAAGATAA 1894
Db 1621 CAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTCATGCTAAAGATAA 1680
QY 1895 AGGAGGCTTTGTACCTTTGCACATGTCATCTTATGGACATATGAAGTTGCAGAACT 1954
Db 1681 AGGAGGCTTTGTACCTTTGCACATGTCATCTTATGGACATATGAAGTTGCAGAACT 1740
QY 1955 TCTTGTAAACATGGAGCAGTAGTAAATAGTGTGATTTATGGAATTTACACCTTTACA 2014
Db 1741 TCTTGTAAACATGGAGCAGTAGTAAATAGTGTGATTTATGGAATTTACACCTTTACA 1800
QY 2015 TGAAGCAGCAGCAAAAGGAAATATGAATTTGCAAACTTTCTGCTCCAGCATGGTCGAGA 2074
Db 1801 TGAAGCAGCAGCAAAAGGAAATATGAATTTGCAAACTTTCTGCTCCAGCATGGTCGAGA 1860
QY 2075 CCCTACCAAAAAACAGGGATGGAATACTCTCTTTGGATCTTTGTTAAAGATGGAGATAC 2134
Db 1861 CCCTACCAAAAAACAGGGATGGAATACTCTCTTTGGATCTTTGTTAAAGATGGAGATAC 1920
QY 2135 AGATATTCAGATCTGCTTAGGGGAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTG 2194
Db 1921 AGATATTCAGATCTGCTTAGGGGAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTG 1980
QY 2195 TTTAGCCAGAGTGAAGAAGTTGCTTCTCTGATATGTAATTTAGAACTTTGCGGATACCCAAGG 2254
Db 1981 TTTAGCCAGAGTGAAGAAGTTGCTTCTCTGATATGTAATTTAGAACTTTGCGGATACCCAAGG 2040
QY 2255 CAGACATTTCAACACCTTTACATTTAGCAGCTGGTTATATAATTTAGAACTTTGCGAGTA 2314
Db 2041 CAGACATTTCAACACCTTTACATTTAGCAGCTGGTTATATAATTTAGAACTTTGCGAGTA 2100
QY 2315 TTTGTTTACACAGGAGCTGATGTAATGCCAAGCAAAAGGAGGACTTAATCTTTTACA 2374
Db 2101 TTTGTTTACACAGGAGCTGATGTAATGCCAAGCAAAAGGAGGACTTAATCTTTTACA 2160
QY 2375 TAATGAGCAGCTTTACGGGATGTAGATGTAGCAGCTCTACTAATAAAGTATATGTCATG 2434
Db 2161 TAATGAGCAGCTTTACGGGATGTAGATGTAGCAGCTCTACTAATAAAGTATATGTCATG 2220
QY 2435 TGTCAATGCCAGGCAAAATGGCTTTACACCTTTGCACGAAGCAGCCCAAAAGGGAGC 2494
Db 2221 TGTCAATGCCAGGCAAAATGGCTTTACACCTTTGCACGAAGCAGCCCAAAAGGGAGC 2280
QY 2495 AACACAGCTTTTGTGCTTTGCTAGCCCATGGAGCTGACCCGACCTTAAAAATCAGGA 2554
Db 2281 AACACAGCTTTTGTGCTTTGCTAGCCCATGGAGCTGACCCGACCTTAAAAATCAGGA 2340
QY 2555 AGGACAAACACCTTTAGTTTGTAGTTTACGGGATGTATGTCAGCGCTTCTTCTGACAGCAGC 2614
Db 2341 AGGACAAACACCTTTAGTTTGTAGTTTACGGGATGTATGTCAGCGCTTCTTCTGACAGCAGC 2400
QY 2615 CATGCCCATCTGCTCTGCCCTCTTTTACAGGCTCAAGTGTCAATGGTGTGAGAG 2674
Db 2401 CATGCCCATCTGCTCTGCCCTCTTTTACAGGCTCAAGTGTCAATGGTGTGAGAG 2460
QY 2675 CCCAGGAGCAGCTGAGATGCTCTCTTTCAGGTCCATCTAGCCCATCAAGCTTCTGTC 2734
Db 2461 CCCAGGAGCAGCTGAGATGCTCTCTTTCAGGTCCATCTAGCCCATCAAGCTTCTGTC 2520
QY 2735 AGCCAGAGCTTTGACAACTTATCTGGGAGTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2794
Db 2521 AGCCAGAGCTTTGACAACTTATCTGGGAGTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2580
QY 2795 AAGTGAACAGAGGCTGCTCCAGTTTGGAGAAAAGAGGTTCCAGGAGTAGATTTAG 2854
Db 2581 AAGTGAACAGAGGCTGCTCCAGTTTGGAGAAAAGAGGTTCCAGGAGTAGATTTAG 2640
QY 2855 CATAACTCAATTCGTAAAGGAATCTTGGACTTGGACACCTAATGGATATATTGAGAGAGA 2914

Db 2641 CATAACTCAATTCGTAAGGAATCTTGGACTTGAACCTTAATGGATATATTGAGAGAGA 2700
QY 2915 ACAGATCACCTTTGATGTATTAGTTGAGATGGGCGCAAAAGGAGCTGAAGGAGATTGGAAT 2974
Db 2701 ACAGATCACCTTTGATGTATTAGTTGAGATGGGCGCAAAAGGAGCTGAAGGAGATTGGAAT 2760
QY 2975 CAATGCTTATGGACATAGGCACAACTAAATTAAGGAGTCCGAGAGCTTATCTCCGGACA 3034
Db 2761 CAATGCTTATGGACATAGGCACAACTAAATTAAGGAGTCCGAGAGCTTATCTCCGGACA 2820
QY 3035 ACAAGGCTTTAAACCATATTTAACTTTGAACACCTCTGGTAGTGGAACTTTCTTATAGA 3094
Db 2821 ACAAGGCTTTAAACCATATTTAACTTTGAACACCTCTGGTAGTGGAACTTTCTTATAGA 2880
QY 3095 TCTGCTCTCTGATGATAAAGAGTTTCACTCTGTGGAGGAAGAGATGCAAAAGTACAGTTCCG 3154
Db 2881 TCTGCTCTCTGATGATAAAGAGTTTCACTCTGTGGAGGAAGAGATGCAAAAGTACAGTTCCG 2940
QY 3155 AGAGCACAGAGATGGAGTTCATGCAAGTGAATCTTCAACAGATACAAATATCTCAAGAT 3214
Db 2941 AGAGCACAGAGATGGAGTTCATGCAAGTGAATCTTCAACAGATACAAATATCTCAAGAT 3000
QY 3215 TCAGAAGGTTTGTAAACAAGAACTATGGGAAAGATACACTCACCGGAGAAAAGATTTTC 3274
Db 3001 TCAGAAGGTTTGTAAACAAGAACTATGGGAAAGATACACTCACCGGAGAAAAGATTTTC 3060
QY 3275 TGAAGAAACCAACCATGCCAATGAACGAATGCTATTTTCATGGGCTCTCTTTTGTGAA 3334
Db 3061 TGAAGAAACCAACCATGCCAATGAACGAATGCTATTTTCATGGGCTCTCTTTTGTGAA 3120
QY 3335 TGCAATTTATCCCAAAAGGCTTTGATGAAAGGCTATGCGTACATAGGTGGTATGTTGGAGC 3394
Db 3121 TGCAATTTATCCCAAAAGGCTTTGATGAAAGGCTATGCGTACATAGGTGGTATGTTGGAGC 3180
QY 3395 TGGCAATTTATTTGCTGAAACCTCTTCCAAAAGCAATCAATATGTATATGTAATTTGGAGG 3454
Db 3181 TGGCAATTTATTTGCTGAAACCTCTTCCAAAAGCAATCAATATGTATATGTAATTTGGAGG 3240
QY 3455 AGTACTCGGTGTCCAGTTTCACAAAGACAGATCTTGTACATTTGCCACAGGAGCTGCT 3514
Db 3241 AGTACTCGGTGTCCAGTTTCACAAAGACAGATCTTGTACATTTGCCACAGGAGCTGCT 3300
QY 3515 CTTTTCGCCGGTAACTTGGGAAAGTCTTTCCTGCAAGTTTCAGTGCATGCAATGAAATGCGACA 3574
Db 3301 CTTTTCGCCGGTAACTTGGGAAAGTCTTTCCTGCAAGTTTCAGTGCATGCAATGAAATGCGACA 3360
QY 3575 TTCTCTCCAGGTCACTCACTGCTAGTGGAGCCAGTGTAAATGGCCTAGCATTAGC 3634
Db 3361 TTCTCTCCAGGTCACTCACTGCTAGTGGAGCCAGTGTAAATGGCCTAGCATTAGC 3420
QY 3635 TGAATATGTTATTTACAGAGGAGACAGGCTTATCTGAGTATTTAACTTACCAGAT 3694
Db 3421 TGAATATGTTATTTACAGAGGAGACAGGCTTATCTGAGTATTTAACTTACCAGAT 3480
QY 3695 TATGAGGCTGAAGGTATGTCGATGGA 3722
Db 3481 TATGAGGCTGAAGGTATGTCGATGGA 3508

RESULT 10

AAA91487

ID AAA91487 standard; DNA; 3797 bp.

XX AAA91487;

AC AAA91487;

XX 06-AUG-2001 (first entry)

DT Tankyrase homologue isotype 1 coding sequence.

DE Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;

XX cell cycle protein; cell cycle associated disorder; cancer; gene mapping;

KW chromosome mapping; gene therapy; vaccine; ds.

KW

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX CDS 404..3706

FT /*tag= a

FT /partial

FT /product= "TaHo1"

FT /note= "TANKYRASE homologue isotype 1; No start

FT codon given"

XX WO200130987-A2.

XX 03-MAY-2001.

XX 25-OCT-2000; 2000WO-US41528.

XX 25-OCT-1999; 99US-0427154.

XX (RIGE-) RIGEL PHARM INC.

XX PA Luo Y, Chan E, Xu X, Huang B;

XX PI WPI; 2001-300503/31.

XX DR P-PSDB; AAY97748.

XX Novel recombinant cell cycle polypeptide, tankyrase H useful for

PT inducing or preventing cell proliferation in cells, and for diagnosing,

PT treating or preventing cell cycle associated disorders such as cancer

PT .

XX Claim 13; Fig 1; 63pp; English.

XX This sequence encodes the Tankyrase homologue isotype 1 (TaHo-1) protein

CC of the invention. The invention also relates to the TaHo-2 protein.

CC The TaHo proteins are useful for inducing or preventing cell

CC proliferation in cells, and in the study or treatment of conditions

CC mediated by the cell cycle proteins, such as to diagnose, treat or

CC prevent cell cycle associated disorders, preferably cancer. The TaHo

CC coding sequences are useful as hybridisation probes, in chromosome and

CC gene mapping and in the generation of anti-sense DNA and RNA. The coding

CC sequences are also useful for the preparation of TaHo, for generating

CC either transgenic animals or knock out animals which, in turn, are useful

CC in a development and screening of therapeutically useful agents, in gene

CC therapy, as vaccine, and for construction of hybridisation probes for

CC mapping the gene which encodes TaHo and for the genetic analysis of

CC individuals with genetic disorders. The TaHo proteins, and their coding

CC sequences are useful in screening assays.

XX SQ Sequence 3797 BP; 1153 A; 732 C; 850 G; 1062 T; 0 other;

Query Match 88.9%; Score 3393.4; DB 22; Length 3797;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 422 AGGTTTGGCGGAAAGAGCTAGTTGAATATTTGCTTCAGAAATGGTGCAATGTCCTCAAGC 481

DB 403 AGGTTTGGCGGAAAGAGCTAGTTGAATATTTGCTTCAGAAATGGTGCAATGTCCTCAAGC 462

QY 482 AGGTGATGATGGGGGCTTATTCCTTCATTAATGCAATGCTCTTTTGGTCATGCTGAAGT 541

DB 463 AGGTGATGATGGGGGCTTATTCCTTCATTAATGCAATGCTCTTTTGGTCATGCTGAAGT 522

QY 542 AGTCAATCTCCTTTTGGACATGGTGCAGACCCCAATGCTGCAGATAATTTGAATTTATAC 601

DB 523 AGTCAATCTCCTTTTGGACATGGTGCAGACCCCAATGCTGCAGATAATTTGAATTTATAC 582

QY 602 TCCTCTCCATGAAGCTGCAATTTAAAGGAAAGATTTGATGTTTGCATTTGCTGTTACAGCA 661

DB 583 TCCTCTCCATGAAGCTGCAATTTAAAGGAAAGATTTGATGTTTGCATTTGCTGTTACAGCA 642

QY 662 TCGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGACATCC 721

DB 643 TGGAGCTGAGCCCAACCATCCGAAATACAGATGAAGGACAGCATTTGGATTTAGCAGATCC 702

QY 722 ATCTGCCAAGCAGTGTCTTACTGGTCAATATAGAAGAGATGAATCTTTAGAAGTCCAG 781

DB 703 ATCTGCCAAGCAGTGTCTTACTGGTCAATATAGAAGAGATGAATCTTTAGAAGTCCAG 762

QY 782 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 841

DB 763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822

QY 842 AAGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATACAGAGTAAAGAT 901

DB 823 AAGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATACAGAGTAAAGAT 882

QY 902 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGTAAAGATAAAGTGAATCTGGT 961

DB 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGTAAAGATAAAGTGAATCTGGT 942

QY 962 ACCATTACACAATGCTGTCTTATGGTCAATATGAAGTAACATGAACCTTTTGGTCAAGCA 1021

DB 943 ACCATTACACAATGCTGTCTTATGGTCAATATGAAGTAACATGAACCTTTTGGTCAAGCA 1002

QY 1022 TGGTGCCTGTGTAATGCAATGGACTTGGGCAATTTCACTCTCTTCATGAGGCACTTC 1081

DB 1003 TGGTGCCTGTGTAATGCAATGGACTTGGGCAATTTCACTCTCTTCATGAGGCACTTC 1062

QY 1082 TAAGAACAGGGTTGAAGTATGTTCTTCTTAAAGTTATGGTGCAGACCCCAACTGCT 1141

DB 1063 TAAGAACAGGGTTGAAGTATGTTCTTCTTAAAGTTATGGTGCAGACCCCAACTGCT 1122

QY 1142 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCACACCACAGTTAAAGAAAGATT 1201

DB 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCACACCACAGTTAAAGAAAGATT 1182

QY 1202 AGCATATGAATTTAAAGGCCACCTGTTGCTGCAAGTGCACGAGAGCTGATCTTACTCG 1261

DB 1183 AGCATATGAATTTAAAGGCCACCTGTTGCTGCAAGTGCACGAGAGCTGATCTTACTCG 1242

QY 1262 AATCAAAAAACATCTCTCTCGGAAATGGTGAATTTCAAGCATCCTCAACACATGAAC 1321

DB 1243 AATCAAAAAACATCTCTCTCGGAAATGGTGAATTTCAAGCATCCTCAACACATGAAC 1302

QY 1322 AGCATTCATTTGCTGCTGCTGCATCTCCATATCCAAAGAAAGCAATATGTCAACTGTT 1381

DB 1303 AGCATTCATTTGCTGCTGCTGCATCTCCATATCCAAAGAAAGCAATATGTCAACTGTT 1362

QY 1382 GCTAAGAAAGAGGCAACATCAATGAAAGACTTAAAGAAATTTCTTGACTCCTCTGCACGT 1441

DB 1363 GCTAAGAAAGAGGCAACATCAATGAAAGACTTAAAGAAATTTCTTGACTCCTCTGCACGT 1422

QY 1442 GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGT 1501

DB 1423 GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGT 1482

QY 1502 TAATGCTCTGATAATCTTGGTCAGACTTCTTACACAGAGCTCCATATTTGGTCACTCT 1561

DB 1483 TAATGCTCTGATAATCTTGGTCAGACTTCTTACACAGAGCTCCATATTTGGTCACTCT 1542

QY 1562 ACAACCTGCGGCTCTCTCTGAGCTATGGGTGTGATCTTAAACATATATCCCTTTCAGGG 1621

DB 1543 ACAACCTGCGGCTCTCTCTGAGCTATGGGTGTGATCTTAAACATATATCCCTTTCAGGG 1602

QY 1622 CTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCCTCAAGAGGATATCTC 1681

DB 1603 CTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCCTCAAGAGGATATCTC 1662

QY 1682 ATTAGGTAATTCAGAGGACAGACACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGA 1741

DB 1663 ATTAGGTAATTCAGAGGACAGACACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGA 1722

QY 1742 AACTGTAAAAAACTGTGTACTGTTCAAGTGTCAACTGCAGACACATTTGAAGGCGCTCA 1801

DB 1723 AACTGTAAAAAACTGTGTACTGTTCAAGTGTCAACTGCAGACACATTTGAAGGCGCTCA 1782

QY	1802	GTCTACACCACCTTCATTTTGCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCT	186
DB	1783	GTCTACACCACCTTCATTTTGCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCT	1842
QY	1862	ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGAGGCCCTTGTAACCTTTGGCACAAATGC	1921
DB	1843	ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGAGGCCCTTGTAACCTTTGGCACAAATGC	1902
QY	1922	ATGTTCTTATGGACATATTAAGACTTGCAGAACTTCTGTTAAACATGGACGAGTAGTTAA	1981
DB	1903	ATGTTCTTATGGACATATTAAGACTTGCAGAACTTCTGTTAAACATGGACGAGTAGTTAA	1962
QY	1982	TGTAGCTGATTTATGGAAATTTACACCTTTTACATGAAGCAGCAGCAAAAGAAAATATGA	2041
DB	1963	TGTAGCTGATTTATGGAAATTTACACCTTTTACATGAAGCAGCAGCAAAAGAAAATATGA	2022
QY	2042	AAATTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACCAAAAAAACAGGGATGGAAA	2101
DB	2023	AAATTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACCAAAAAAACAGGGATGGAAA	2082
QY	2102	TACTCCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA	2161
DB	2083	TACTCCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA	2142
QY	2162	TGCAGCTTGTCTAGATGCTGCCAAGAGGGTTGTTTATGCCAGAGTGGAAGATGTCTTTC	2221
DB	2143	TGCAGCTTGTCTAGATGCTGCCAAGAGGGTTGTTTATGCCAGAGTGGAAGATGTCTTTC	2202
QY	2222	TCCTGTATATGTAAATTGCCGGGATACCCAAGGAGACATTCACACCTTTACATTTAGC	2281
DB	2203	TCCTGTATATGTAAATTGCCGGGATACCCAAGGAGACATTCACACCTTTACATTTAGC	2262
QY	2282	AGCTGGTTATATAATTTAGAAGTTGCAGAGTATTTGTTACACACGAGGAGCTGATGTGA	2341
DB	2263	AGCTGGTTATATAATTTAGAAGTTGCAGAGTATTTGTTACACACGAGGAGCTGATGTGA	2322
QY	2342	TGCCCAAGACAAGGAGGACTTATTCCCTTTACATAAATGCAGCATCTACGGGCATGTAGA	2401
DB	2323	TGCCCAAGACAAGGAGGACTTATTCCCTTTACATAAATGCAGCATCTACGGGCATGTAGA	2382
QY	2402	TGTAGAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGAGCAAAATGGGCTTT	2461
DB	2383	TGTAGAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGAGCAAAATGGGCTTT	2442
QY	2462	CACACCTTTGCAGAGCAGCCCAAAAGGGAGCAACACAGCTTTGTGCTTGTGCTAGC	2521
DB	2443	CACACCTTTGCAGAGCAGCCCAAAAGGGAGCAACACAGCTTTGTGCTTGTGCTAGC	2502
QY	2522	CCATGAGCTGACCCGACTCTTAAAAATCAGGAAGCAACACCTTTAGATTAGTTTTC	2581
DB	2503	CCATGAGCTGACCCGACTCTTAAAAATCAGGAAGCAACACCTTTAGATTAGTTTTC	2562
QY	2582	AGCGGATGATGCAGGCTCTCTGCAGAGCACCATGCCCCCATCTGCTCGCCCTCTTG	2641
DB	2563	AGCGGATGATGCAGGCTCTCTGCAGAGCACCATGCCCCCATCTGCTCGCCCTCTTG	2622
QY	2642	TTACAAGCCTCAAGTGTCTCAATGTGTGAGAAGCCCAAGAGCCACTGCAGATGCTCTTC	2701
DB	2623	TTACAAGCCTCAAGTGTCTCAATGTGTGAGAAGCCCAAGAGCCACTGCAGATGCTCTTC	2682
QY	2702	TTCAAGTCCATCTAGCCCATCAAGCCCTTTCTGCAGCCAGCAGTCTTGACAACCTTATCTGG	2761
DB	2683	TTCAAGTCCATCTAGCCCATCAAGCCCTTTCTGCAGCCAGCAGTCTTGACAACCTTATCTGG	2742
QY	2762	GAGTTTTCAGAACTTGCTTTCAGTAGTTAGTTTCAAGTGAAGAGGGTGTCTTCCAGTTT	2821
DB	2743	GAGTTTTCAGAACTTGCTTTCAGTAGTTAGTTTCAAGTGAAGAGGGTGTCTTCCAGTTT	2802
QY	2822	GGAGAAAAGAGGTTTCAGAGTAGATTTTATAGATAACTCAATTCGTAAAGGAATCTTGG	2881
DB	2803	GGAGAAAAGAGGTTTCAGAGTAGATTTTATAGATAACTCAATTCGTAAAGGAATCTTGG	2862

QY	2882	AC	TTGAGCACCCTAATGGATATATTTCGAGAGAACAGATCACTTTGATGTATTAGTTGA	2941
DB	2863	AC	TTGAGCACCCTAATGGATATATTTCGAGAGAACAGATCACTTTGATGTATTAGTTGA	2932
QY	2942	GAT	GGGCAACAAGAGCTGAAGAGATTGGAATCAATGCTTATGGACATAGGCACAACCT	3001
DB	2923	GAT	GGGCAACAAGAGCTGAAGAGATTGGAATCAATGCTTATGGACATAGGCACAACCT	2982
QY	3002	AA	TTAAAGAGAGTCGAGAGACTTATCTCCGACACAACAGGTCCTTAACCCATATTAACTTT	3061
DB	2983	AA	TTAAAGAGAGTCGAGAGACTTATCTCCGACACAACAGGTCCTTAACCCATATTAACTTT	3042
QY	3062	GA	ACACCTCGTAGTGGGAACAATCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA	3121
DB	3043	GA	ACACCTCGTAGTGGGAACAATCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA	3102
QY	3122	GT	CTGTGGAGGAAGAGATGCAAAAGTACAGTTTCAGAGCACAGAGATGGAGGTCATGCAGG	3181
DB	3103	GT	CTGTGGAGGAAGAGATGCAAAAGTACAGTTTCAGAGCACAGAGATGGAGGTCATGCAGG	3162
QY	3182	TG	GAATCTTCAACAGATACAATATCTCAAGATTTCAGAGGTTTGTAAACAAGAAACTATG	3241
DB	3163	TG	GAATCTTCAACAGATACAATATCTCAAGATTTCAGAGGTTTGTAAACAAGAAACTATG	3222
QY	3242	GG	AAGATACACTACCCGGAGAAAGAAGTTTCTGAAGAAACCAACCATGCCAATGA	3301
DB	3223	GG	AAGATACACTACCCGGAGAAAGAAGTTTCTGAAGAAACCAACCATGCCAATGA	3282
QY	3302	AC	GAATGCTATTTCATGGGCTCTCCTTTTGTGAATGCAATTATCCACAAGGCTTTGATGA	3361
DB	3283	AC	GAATGCTATTTCATGGGCTCTCCTTTTGTGAATGCAATTATCCACAAGGCTTTGATGA	3342
QY	3362	AA	GGCATGCGTACATAGTGGTATGTTTGGAGCTGGCATTATTTTTGTCTCAAAACTCTTC	3421
DB	3343	AA	GGCATGCGTACATAGTGGTATGTTTGGAGCTGGCATTATTTTTGTCTCAAAACTCTTC	3402
QY	3422	CA	AAAGCAATCAATATGTATATGGAATTGGAGAGAGTACTGGGTGCCAGTTTCAACAAGA	3481
DB	3403	CA	AAAGCAATCAATATGTATATGGAATTGGAGAGAGTACTGGGTGCCAGTTTCAACAAGA	3462
QY	3482	CAG	ACTTGTACATTGGCCACAGCAGCTGCTCTTTTCCCGGGTAACCTTGGGAAAGTC	3541
DB	3463	CAG	ACTTGTGTACATTGGCCACAGCAGCTGCTCTTTTCCCGGGTAACCTTGGGAAAGTC	3522
QY	3542	TT	CTCTGCAGTTTCAGTGCATCAAAATGGCAATCTCTCTCCAGGTCATCACTTCAGTCAC	3601
DB	3523	TT	CTCTGCAGTTTCAGTGCATCAAAATGGCAATCTCTCTCCAGGTCATCACTTCAGTCAC	3582
QY	3602	TGG	TAGGCCAGTGTAATGGCCTAGCAATTAGCTGAATATGTTATTACAGAGGAGAACA	3661
DB	3583	TG	TAGGCCAGTGTAATGGCCTAGCAATTAGCTGAATATGTTATTACAGAGGAGAACA	3642
QY	3662	GG	CTTATCTTGAGTATTAAATTACATTACCAGATTATGAGGCTGAAGGTATGGTCGATGG	3721
DB	3643	GG	CTTATCTTGAGTATTAAATTACATTACCAGATTATGAGGCTGAAGGTATGGTCGATGG	3702
QY	3722	AT	AAATAGTATTTTAAGAAACTAATTCACGTGAACCTAAAATCATCAAGCAGCAGTCG	3781
DB	3703	AT	AAATAGTATTTTAAGAAACTAATTCACGTGAACCTAAAATCATCAAGCAGCAGTCG	3762
QY	3782	CT	CTACGCTTTTACCTCTTTGCTGAAAAAATAAAAA 3816	
DB	3763	CT	CTACGCTTTTACCTCTTTGCTGAAAAAATAAAAA 3797	

RESULT 11
AAC66824
ID AAC66824 standard; cDNA; 4296 BP.
XX
XX AAC66824;
XX
XX
DT 27-FEB-2001 (first entry)
XX

QY 741 ACTGGTGAATATAGAAGATGAACCTTAGAAAGTCCAGGAGTGGCAATGAAGAAAA 800
Db 241 ACTGGTGAATATAGAAGATGAACCTTAGAAAGTCCAGGAGTGGCAATGAAGAAAA 300
QY 801 ATGATGGCTCTACTACACCAATTAATATGCACTGCCAGCAAGTATGGCAGAAAAAGTCA 860
Db 301 ATGATGGCTCTACTACACCAATTAATATGCACTGCCAGCAAGTATGGCAGAAAAAGTCA 360
QY 861 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGATTGTACAGCTGTACTGCA 920
Db 361 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGATTGTACAGCTGTACTGCA 420
QY 921 CATGGAGCTGATGTCATGCTAAAGATAAAGGTGATCTGGTACCAATACCAATGCCTGT 980
Db 421 CATGGAGCTGATGTCATGCTAAAGATAAAGGTGATCTGGTACCAATACCAATGCCTGT 480
QY 981 TCTATGGTCAATATGAAGTAACTGAACCTTTGGTCAAGCATGGTGCCTGTGTAATGCA 1040
Db 481 TCTATGGTCAATATGAAGTAACTGAACCTTTGGTCAAGCATGGTGCCTGTGTAATGCA 540
QY 1041 ATGGACTTTGGCAATTCACCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 1100
Db 541 ATGGACTTTGGCAATTCACCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 600
QY 1101 TGTCTCTCTCTTAAGTTATGGTCAGACCCAAACACTGCTCAATTTGCACAAATAAAGT 1160
Db 601 TGTCTCTCTCTTAAGTTATGGTCAGACCCAAACACTGCTCAATTTGAAGATAAAGT 660
QY 1161 GCTATAGACTTGGCTCCCAACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAGGC 1220
Db 661 GCTATAGACTTGGCTCCCAACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAGGC 720
QY 1221 CACTCGTGTCTGCAAGCTGCACGAGAAGCTGATGTACTCGAATCAAAAACATCTCTCT 1280
Db 721 CACTCGTGTCTGCAAGCTGCACGAGAAGCTGATGTACTCGAATCAAAAACATCTCTCT 780
QY 1281 CTGGAATAGGTCAATTTCAAGCATCTCAACACATGAACAGATGTCATTTGCTGTCT 1340
Db 781 CTGGAATAGGTCAATTTCAAGCATCTCAACACATGAACAGATGTCATTTGCTGTCT 840
QY 1341 GCATCTCCATATCCCAAGAAAGCAATATATGTAACCTTTGCTAAGAAAGGACCAAC 1400
Db 841 GCATCTCCATATCCCAAGAAAGCAATATATGTAACCTTTGCTAAGAAAGGACCAAC 900
QY 1401 ATCAATGAAGACATTAAGAAATTTTGACTCTCTGACGTGGCATCTGAGAAAGCTCAT 1460
Db 901 ATCAATGAAGACATTAAGAAATTTTGACTCTCTGACGTGGCATCTGAGAAAGCTCAT 960
QY 1461 AATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTGGATATCTT 1520
Db 961 AATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTGGATATCTT 1020
QY 1521 GGTACAGCTTCTTACAGAGCTGCATATTTGTTGTCATCTACAAACCTGCCGCTACTC 1580
Db 1021 GGTACAGCTTCTTACAGAGCTGCATATTTGTTGTCATCTACAAACCTGCCGCTACTC 1080
QY 1581 CTGAGCTATGGTGTGATCTTAACATTTATCCCTTCAGGCTTTCACCTTACAGATG 1640
Db 1081 CTGAGCTATGGTGTGATCTTAACATTTATCCCTTCAGGCTTTCACCTTACAGATG 1140
QY 1641 GGAATGAATATGATACAGCAACTCTCTCAAGAGGATATCTCATTTAGSTAAATTCAGAGCA 1700
Db 1141 GGAATGAATATGATACAGCAACTCTCTCAAGAGGATATCTCATTTAGSTAAATTCAGAGCA 1200
QY 1701 GACAGCAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGAAACTGTAAAAAACTGTGT 1760
Db 1201 GACAGCAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGAAACTGTAAAAAACTGTGT 1260
QY 1761 ACTGTTACAGTGTCAACTGCAGAGACATTTGAAGGGCTGAGTCTACACCCTCATTTT 1820
Db 1261 ACTGTTACAGTGTCAACTGCAGAGACATTTGAAGGGCTGAGTCTACACCCTCATTTT 1320
QY 1821 GCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTCATGTG 1880

Db 1321 GCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGAGCTGATGTG 1380
QY 1881 CAPGCTAAAGATAAAGAGGAGGCTTTGACCTTTGCACAATGCATGTTCTTTATGGACATTAT 1940
Db 1381 CATGCTAAAGATAAAGAGGAGGCTTTGACCTTTGCACAATGCATGTTCTTTACGGACATTAT 1440
QY 1941 GAAGTTGCAGAACTTCTTGTAAACATGAGCAGTAGTTAATGTAGCTGATTTATGAAA 2000
Db 1441 GAAGTTGCAGAACTTCTTGTAAACATGAGCAGTAGTTAATGTAGCTGATTTATGAAA 1500
QY 2001 TTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAAAATATGAAAATTTGCAAACTTCTGCTC 2060
Db 1501 TTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAAAATATGAAAATTTGCAAACTTCTGCTC 1560
QY 2061 CAGCATGTGTCAGACCTTACCACCAAAAAACAGGGATGGAATACTCCTTTGGATCTTGT 2120
Db 1561 CAGCATGTGTCAGACCTTACCACCAAAAAACAGGGATGGAATACTCCTTTGGATCTTGT 1620
QY 2121 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 2180
Db 1621 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 1680
QY 2181 GCCAAGAGGGTGTGTAGCCAGAGTGAAGAAAGTTGCTTCTCCTGATATGTAATTTGC 2240
Db 1681 GCCAAGAGGGTGTGTAGCCAGAGTGAAGAAAGTTGCTTCTCCTGATATGTAATTTGC 1740
QY 2241 CGCGATACCAAGGAGCAGATTTCAACACCTTTACATTTAGCAGCTGGTTATATAATTTA 2300
Db 1741 CGCGATACCAAGGAGCAGATTTCAACACCTTTACATTTAGCAGCTGGTTATATAATTTA 1800
QY 2301 GAAGTTGCAGACTATTGTTACAACACGAGCTGATGTAATGCCCAACACAAAGGAGGA 2360
Db 1801 GAAGTTGCAGACTATTGTTACAACACGAGCTGATGTAATGCCCAACACAAAGGAGGA 1860
QY 2361 CTTATTCCTTTTACATTAATGACAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2420
Db 1861 CTTATTCCTTTTACATTAATGACAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920
QY 2421 AAGTATATGTCATGTCAATGCCAGGACAAATGGGCTTTTACACCTTTGACAGCAAGA 2480
Db 1921 AAGTATATGTCATGTCAATGCCAGGACAAATGGGCTTTTACACCTTTGACAGCAAGA 1980
QY 2481 GCCCAAGAGGACGACACAGCTTTTGTGCTTTGTAGCCCATGGAGCTCACCCGACT 2540
Db 1981 GCCCAAGAGGACGACACAGCTTTTGTGCTTTGTAGCCCATGGAGCTCACCCGACT 2040
QY 2541 CTTTAAAAATCAGGAAGGACAAACACCTTTTAGATTAGTTTTCAGCGGATGATCTCAGCGCT 2600
Db 2041 CTTTAAAAATCAGGAAGGACAAACACCTTTTAGATTAGTTTTCAGCGGATGATCTCAGCGCT 2100
QY 2601 CTTCTGACAGAGCCATGCCCCATCTGCTGCCCTCTTGTTCACAGCCTCAAGTGCTC 2660
Db 2101 CTTCTGACAGAGCCATGCCCCATCTGCTGCCCTCTTGTTCACAGCCTCAAGTGCTC 2160
QY 2661 AATGTTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTCTCAGGTCATCTAGCCCA 2720
Db 2161 AATGTTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTCTCAGGTCATCTAGCCCA 2220
QY 2721 TCAAGCCTTTCTGACGAGCAGCTTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2780
Db 2221 TCAAGCCTTTCTGACGAGCAGCTTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2280
QY 2781 TCAGTAGTTAGTTCAAGTGGAAACAGAGGCTCTCCACTTTGGAGAAAAAGAGGTTCCA 2840
Db 2281 TCAGTAGTTAGTTCAAGTGGAAACAGAGGCTCTCCACTTTGGAGAAAAAGAGGTTCCA 2340
QY 2841 GGAGTAGATTTTACGATAACTCAATTCGTAAGGAATCTTGAGCTTTGAGCAGCTAATGAT 2900
Db 2341 GGAGTAGATTTTACGATAACTCAATTCGTAAGGAATCTTGAGCTTTGAGCAGCTAATGAT 2400
QY 2901 ATATTTGAGAGAGAACACATCCTTTGGATGTATGTAGTTGAGATGGGGCACAAGAGCTG 2960
Db 2901 ATATTTGAGAGAGAACACATCCTTTGGATGTATGTAGTTGAGATGGGGCACAAGAGCTG 2960

Db 2401 ATATTGTGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGCAACAAGGAGCTG 2460
QY 2961 AAGGAGATTGGAATCAATGCTTATGACATAGGCACAACCAATTAATAAGGAGTGCAGAGA 3020
Db 2461 AAGGAGATTGGAATCAATGCTTATGACATAGGCACAACCAATTAATAAGGAGTGCAGAGA 2520
QY 3021 CTTATCTCCGGACAACAAGGCTTTAAACCATATTTAACTTTGAACACCTCTGGTAGTGA 3080
Db 2521 CTTATCTCCGGACAACAAGGCTTTAAACCATATTTAACTTTGAACACCTCTGGTAGTGA 2580
QY 3081 ACAATCTTATAGATCTGCTCCGTCATGATCAATGAAGATTTCACTGTCGGAGAGAGATG 3140
Db 2581 ACAATCTTATAGATCTGCTCCGTCATGATCAATGAAGATTTCACTGTCGGAGAGAGATG 2640
QY 3141 CAAAGTACAGTTTCGAGAGCACAGATGGAGTCAATGAGGTGGAATCTTCAACAGATAC 3200
Db 2641 CAAAGTACAGTTTCGAGAGCACAGATGGAGTCAATGAGGTGGAATCTTCAACAGATAC 2700
QY 3201 AATATCTCAAGATTCAGAAGGTTTGTAAACAAGAACTATGGAAAGATACACTCACCGG 3260
Db 2701 AATATCTCAAGATTCAGAAGGTTTGTAAACAAGAACTATGGAAAGATACACTCACCGG 2760
QY 3261 AGAAGAAGATTTCTGAAGAAACCAACCAATGCCAATGAACGAATGCTATTTTCATGGG 3320
Db 2761 AGAAGAAGATTTCTGAAGAAACCAACCAATGCCAATGAACGAATGCTATTTTCATGGG 2820
QY 3321 TCTCCTTTTGTCAATGCAATTCACAAAGGCTTTGTAAAGGCATGCGTACATAGGT 3380
Db 2821 TCTCCTTTTGTCAATGCAATTCACAAAGGCTTTGTAAAGGCATGCGTACATAGGT 2880
QY 3381 GGTATGTTTGGAGCTGGCATTTATTTTGTGTAATACTCTCCAAAGCAATCAATATGTA 3440
Db 2881 GGTATGTTTGGAGCTGGCATTTATTTTGTGTAATACTCTCCAAAGCAATCAATATGTA 2940
QY 3441 TATGGAATTTGGAGGAGTACTGGGTGTCAGTTTCAAGAGCAGATCTTGTACATTTGC 3500
Db 2941 TATGGAATTTGGAGGAGTACTGGGTGTCAGTTTCAAGAGCAGATCTTGTACATTTGC 3000
QY 3501 CACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCTTCTCCAGTTCAGTGCA 3560
Db 3001 CACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCTTCTCCAGTTCAGTGCA 3060
QY 3561 ATGAAATGGCAATTTCTCCAGTTCATCACTCAGTCACTGGTAGGCCAGCTTAAT 3620
Db 3061 ATGAAATGGCAATTTCTCCAGTTCATCACTCAGTCACTGGTAGGCCAGCTTAAT 3120
QY 3621 GGCCTAGCATTTAGTGAATATGTTATTTACAGAGGAGACAGGCTTATCCTGAGTATTA 3680
Db 3121 GGCCTAGCATTTAGTGAATATGTTATTTACAGAGGAGACAGGCTTATCCTGAGTATTA 3180
QY 3681 ATTACTTACAGATTATGAGGCTGAAGGTATGGTGGATATAATAGTTATTTTAAAGA 3740
Db 3181 ATTACTTACAGATTATGAGGCTGAAGGTATGGTGGATATAATAGTTATTTTAAAGA 3240
QY 3741 AACTAATTCCTGAACTTAAATCATCAAGCAGCAGTGGCCTCTAGCTTTTACTCCTT 3800
Db 3241 AACTAATTCCTGAACTTAAATCATCAAGCAGCAGTGGCCTCTAGCTTTTACTCCTT 3300
QY 3801 TGCTGAAAAAAA 3813
Db 3301 TGCTGAAAAAAA 3313

RESULT 13

AAC66823 standard: cDNA; 4493 BP.

AC XX

AAC66823;

XX XX

27-FEB-2001 (first entry)

XX XX

DE Human tankyrase II coding sequence SEQ ID NO: 1.

XX

Human: tankyrase II; telomere length; signal transduction; ss.
Homo sapiens.
Key Location/Qualifiers
CDS 1..4002
FT /tag= a
FT /transl_except= (pos:595..597,aa:Xaa)
FT /transl_except= (pos:625..630,aa:XaaXaa)
FT /transl_except= (pos:637..639,aa:Xaa)
FT /transl_except= (pos:694..696,aa:Xaa)
FT /transl_except= (pos:724..726,aa:Xaa)
FT /transl_except= (pos:730..732,aa:Xaa)
FT /transl_except= (pos:763..765,aa:Xaa)
FT /transl_except= (pos:790..792,aa:Xaa)
FT /transl_except= (pos:2176..2178,aa:Xaa)
FT /transl_except= (pos:2380..2382,aa:Xaa)
FT /transl_except= (pos:2398..2400,aa:Xaa)
FT /transl_except= (pos:2464..2469,aa:XaaXaa)
FT /transl_except= (pos:3163..3165,aa:Xaa)
FT /note= "Xaa=unknown"
FT /product= "tankyrase II"
FT /partial
XX
PN WO200061813-A1.
XX
PD 19-OCT-2000.
XX
PF 10-APR-2000; 2000WO-US09558.
XX
PR 09-APR-1999; 99US-0128577.
PR 13-APR-1999; 99US-0129123.
XX
PA (GERO-) GERON CORP.
XX
XX Morin GB, Funk WD, Piatyszek MA;
PI WPI; 2000-679503/66.
DR P-PSDB; AAB27209.
XX
XX
PT Novel mammalian Tankyrase II polypeptide and the polynucleotide
PT encoding the polypeptide useful for modulating or maintaining telomere
PT length, replicative capacity, apoptosis, chromosome packing or gene
PT expression -
XX
PS Claim 1; Fig 2; 52pp; English.
XX
CC The present sequence is a version of the human tankyrase II coding
CC sequence. Its protein is thought to be involved in signal transduction in
CC the cell, and to have binding activity for other telomere-associated
CC proteins. It is possible that it plays a role in the regulation of
CC telomere length, thus affecting the replicative ability of the cell. The
CC protein is useful for ribosylating target proteins, for determining
CC tankyrase II binding activity in a sample, and for modulating telomere
CC length in a cell.
XX
SQ Sequence 4493 BP; 1119 A; 685 C; 760 G; 985 T; 944 other;
Query Match 79.6%; Score 3039; DB 21; Length 4493;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 3101; Conservative 1; Mismatches 375; Indels 27; Gaps 2;
QY 337 AACGAGTCAAGAGCTGTGACGCTGAGAAGGTGAACAGCGCGACACGCGGCGAGGA 396
Db 587 AACCATTCNNAGGCTGTGACCCCTGAARAGGTNACANCCNCAACACGCGGCGAGGA 646
QY 397 AATCCACCCCGCTGCACCTCCCGCAGGTTTGGCGGAAAGACGTAGTTGAATATTGC 456
Db 647 AATCCACCCCGCTGCACCTCCCGCAGGTTTGGCGGAAAGACCTNNNTAAATATTGC 706
QY 457 TTCAGATGGTGCAGAAATGTCCAGCACGTGATGGGCGCTTATTCCTCTTCATATG 516
Db 707 TTCAAAATGGTGCAGAAATNTCCAAACACTTTATAATATGGGCGCTTATTCCTCTTCATATG 766

Db 2278 AAGAAACCAACACCATGCCAATGAACAAATGCTATTTCATGGGTCTCCTTTGTGAATG 2337
Qy 3337 CAATTATCCACAAAGGCTTTGATGAAGGCAATGCGTACATAGGTGATGTTGGAGCTG 3396
Db 2338 CAATTATCCACAAAGGCTTTGATGAAGGCAATGCGTACATAGGTGATGTTGGAGCTG 2397
Qy 3397 GCATTATTTTGGCTGAAACCTTTCCAAAGCAATCAATATGATGGAATGGAGGAG 3456
Db 2398 GCATTATTTTGGCTGAAACCTTTCCAAAGCAATCAATATGATGGAATGGAGGAG 2457
Qy 3457 GTACTGGGTGCCAGTTTCCAAAGACAGATCTTGTTCATATTTGCCACAGGAGCTGCTCT 3516
Db 2458 GTACTGGGTGCCAGTTTCCAAAGACAGATCTTGTTCATATTTGCCACAGGAGCTGCTCT 2517
Qy 3517 TTTGCCGGGTACCTTTGGAAAGCTTTTCCCTGCGAGTTTCAGTGCATGAATGAATGSCACATT 3576
Db 2518 TTTGCCGGGTAACTTTGGAAAGCTTTTCCCTGCGAGTTTCAGTGCATGAATGAATGSCACATT 2577
Qy 3577 CTCTCCAGGTCACTCACPCAGTCACTGGTAGGCCAGTGTAAATGGGCTAGCATTAGCTG 3636
Db 2578 CTCTCCAGGTCACTCACPCAGTCACTGGTAGGCCAGTGTAAATGGGCTAGCATTAGCTG 2637
Qy 3637 AATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTAATTACTTACCAGATTA 3696
Db 2638 AATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTAATTACTTACCAGATTA 2697
Qy 3697 TGAGGCCCTGAAGGTATGTCGATGATGAATAATAGTATTTTAAAGAACTAAATCCACTGAA 3756
Db 2698 TGAGGCCCTGAAGGTATGTCGATGATGAATAATAGTATTTTAAAGAACTAAATCCACTGAA 2757
Qy 3757 CCTAAATCATCAAGAGCAGCTAGGCTCTAGCTTTTACTCTTTGCTGAAAAAAA 3813
Db 2758 CCTAAATCATCAAGAGCAGCTAGGCTCTAGCTTTTACTCTTTGCTGAAAAAAA 2814

RESULT 15

AAF63919
ID AAF63919 standard; DNA; 2971 BP.
XX AAF63919;
XX 05-APR-2001 (first entry)
XX Human tankyrase2 clone consensus SEQ ID NO: 92.
XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
XX Inflammatory disorder; ds.
XX Homo sapiens.
XX WO200100849-A1.
XX 04-JAN-2001.
XX 28-JUN-2000; 2000WO-US17827.
XX 29-JUN-1999; 99US-0141582.
XX (ICOS-) ICOS CORP.
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX WPI; 2001-102896/11.
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
XX poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
XX inflammatory and autoimmune disorders -
XX Example 1; Page 152-153; 242pp; English.
XX The present invention provides the protein and coding sequence for the
XX human tankyrase2 protein. This is found in two different versions,
XX designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has

CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to aging and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
SQ Sequence 2971 BP; 915 A; 555 C; 642 G; 859 T; 0 other;
Query Match 63.9%; Score 2439.8; DB 22; Length 2971;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1371 TGTGAACCTGTTGCTAAGAAAGGAGCAACATCAATGAAGAACTAAAGAAATTCCTGACT 1430
Db 1 TGTGAACCTGTTGCTAAGAAAGGAGCAACATCAATGAAGAACTAAAGAAATTCCTGACT 60
Qy 1431 CCTCTGCACGTGCATCTGAGAAGCTCATAATGATGTTGTTCAAGTAGTGGTGAACAT 1490
Db 61 CCTCTGCACGTGCATCTGAGAAGCTCATAATGATGTTGTTCAAGTAGTGGTGAACAT 120
Qy 1491 GAAGCAAAGGTTAATGCTCTGGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1550
Db 121 GAAGCAAAGGTTAATGCTCTGGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180
Qy 1551 TGTGGTCACTACAAACCTGCCGCCCTACTCCTGAGCTATGGGTGTGATCCTTAACATATA 1610
Db 181 TGTGGTCACTACAAACCTGCCGCCCTACTCCTGAGCTATGGGTGTGATCCTTAACATATA 240
Qy 1611 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCCTCCAA 1670
Db 241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCCTCCAA 300
Qy 1671 GAGGGTATCTCATTAGGTAATTCAGAGGCGACAGACAAATTCGTGGAAGCTGCAAGAGCT 1730
Db 301 GAGGGTATCTCATTAGGTAATTCAGAGGCGACAGACAAATTCGTGGAAGCTGCAAGAGCT 360
Qy 1731 GGAGATGTCGAAACCTGTAAGAAACCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACTT 1790
Db 361 GGAGATGTCGAAACCTGTAAGAAACCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACTT 420
Qy 1791 GAAGGGCTGAGTCTACACCACTTCATTTGCGAGCTGGGTATACAGAGTGTCCGTGGTG 1850
Db 421 GAAGGGCTGAGTCTACACCACTTCATTTGCGAGCTGGGTATACAGAGTGTCCGTGGTG 480
Qy 1851 GAATATCTGCTACAGCATGGAGCTGATGTCATGCTTAAAGATAAAGGAGGCTTGTACCT 1910
Db 481 GAATATCTGCTACAGCATGGAGCTGATGTCATGCTTAAAGATAAAGGAGGCTTGTACCT 540
Qy 1911 TTGCACAATGCATGTTCTTATGGACATATGAAGTTCGAGAACTTCTTTTAAACATGGA 1970
Db 541 TTGCACAATGCATGTTCTTATGGACATATGAAGTTCGAGAACTTCTTTTAAACATGGA 600
Qy 1971 GCAGTAGTTATGCTAGCTGATTTATGAAATTTTACACCTTTTACATGAAGCAGCAGCAAAA 2030
Db 601 GCAGTAGTTATGCTAGCTGATTTATGAAATTTTACACCTTTTACATGAAGCAGCAGCAAAA 660
Qy 2031 GGAAATATGAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACAAAAAANAAC 2090
Db 661 GGAAATATGAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACAAAAAANAAC 720
Qy 2091 AGGGATGAAATATCTCCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTCAGATCTG 2150
Db 721 AGGGATGAAATATCTCCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTCAGATCTG 780
Qy 2151 CTTAGGGGAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTCCTTTAGCCAGAGTGAAG 2210
Db 781 CTTAGGGGAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTCCTTTAGCCAGAGTGAAG 840
Qy 2211 AAGTTGCTCTCTCTGATAATGTAATTTGCCGCGATACCCCAAGGCGACATTCACACCT 2270
Db 841 AAGTTGCTCTCTCTGATAATGTAATTTGCCGCGATACCCCAAGGCGACATTCACACCT 900
Qy 2271 TTACATTTAGCAGCTGGTGTATAAATTTAGAGTTGCGAGAGTATTTGTTTACAACACGGA 2330

|||||
Db 248 TATTTGGTTTCCAGAAATGTCCTCAAGCAGCTGATGATGGGGCTTTATTCCTCTT 307
QY 510 CATAAATGCATGCTCTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGGCAGCATGGTGCA 569
Db 308 CATAAATGCATGCTCTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGGCAGCATGGTGCA 367
QY 570 GACCCCAATGCTCGAGATAATTTGGAATATATACCTCTCTCCATGAAGCTGCAATTTAAAGGA 629
Db 368 GACCCCAATGCTCGAGATAATTTGGAATATATACCTCTCTCCATGAAGCTGCAATTTAAAGGA 427
QY 630 AAGATTGATGTTCCATGTTGCTGTTACAGCATGGAGCTGAGCCAAACCTCCGAAATPACA 689
Db 428 AAGATTGATGTTGCTGTTGCTGTTACAGCATGGAGCTGAGCCAAACCTCCGAAATPACA 487
QY 690 GATGGAAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAAGCAGTGTACTGGTGAA 749
Db 488 GATGGAAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAAGCAGTGTACTGGTGAA 547
QY 750 TATAAGAAAGATGAACCTTTAGAAAGTGCAGGAGTGGCAATGAAGAAATAATGATGGCT 809
Db 548 TATAAGAAAGATGAACCTTTAGAAAGTGCAGGAGTGGCAATGAAGAAATAATGATGGCT 607
QY 810 CTACTCACACATTAATGTCACAGTAAAGATTGTACAGCTGTTACTGCAACATGAGCT 869
Db 608 CTACTCACACATTAATGTCACAGTAAAGATTGTACAGCTGTTACTGCAACATGAGCT 667
QY 870 CATTTGGCAGCAGGATATAACAGATTAAGATTGTACAGCTGTTACTGCAACATGAGCT 929
Db 668 CATTTGGCAGCAGGATATAACAGATTAAGATTGTACAGCTGTTACTGCAACATGAGCT 727
QY 930 GATGTCATGCTTAAGATAAAGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGT 989
Db 728 GATGTCATGCTTAAGATAAAGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGT 787
QY 990 CATTATGAAGTAACTGAACCTTTTGGTCAAGCATGGTGCCTGTAAATGCAATGGAGCTTG 1049
Db 788 CATTATGAAGTAACTGAACCTTTTGGTCAAGCATGGTGCCTGTAAATGCAATGGAGCTTG 847
QY 1050 TGGCAATTAACCTCTTTCATGAGGAGCTTCTTAAGAACAGGGTTGAAGTATGTTCTCTT 1109
Db 848 TGGCAATTAACCTCTTTCATGAGGAGCTTCTTAAGAACAGGGTTGAAGTATGTTCTCTT 907
QY 1110 CTCTTAAGTATGCTGCAGACCCCAACACTGCTCAATTTGTCACAAATAAAGTGTATAGAC 1169
Db 908 CTCTTAAGTATGCTGCAGACCCCAACACTGCTCAATTTGTCACAAATAAAGTGTATAGAC 967
QY 1170 TTGGCTCCACACACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1229
Db 968 TTGGCTCCACACACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1027
QY 1230 CTGCAAGCTGCAGGAGAGCTGATGTTACTCGAATCAAAAACATCTCTCTGGAATG 1289
Db 1028 CTGCAAGCTGCAGGAGAGCTGATGTTACTCGAATCAAAAACATCTCTCTGGAATG 1087
QY 1290 GTGAATTTCAAGCATCTCAACACATGAACAGCATTTGCAATGCTGTCATCTCA 1349
Db 1088 GTGAATTTCAAGCATCTCAACACATGAACAGCATTTGCAATGCTGTCATCTCA 1147
QY 1350 TATCCCAAAAGAAAGCAATATGTGAACCTGTTGCTAAGAAAGGAGCAACATCAATGAA 1409
Db 1148 TATCCCAAAAGAAAGCAATATGTGAACCTGTTGCTAAGAAAGGAGCAACATCAATGAA 1207
QY 1410 AAGACTAAAGAAATTTTGACTCTCTGACGTTGGCATCTGAGAAAGCTCATATGATGT 1469
Db 1208 AAGACTAAAGAAATTTTGACTCTCTGACGTTGGCATCTGAGAAAGCTCATATGATGT 1267
QY 1470 GTTGAAGTAGTGTGAACATGAACAAAGGTTAATGCTCTGATAATCTTGGTCAGACT 1529
Db 1268 GTTGAAGTAGTGTGAACATGAACAAAGGTTAATGCTCTGATAATCTTGGTCAGACT 1327
QY 1530 TCTCTACAGAGCTGCATATTTGGTGTCTATCTACAAACCTGCGGCTACTCTGAGCTAT 1589
|||||

Db 1328 TCTCTACACAGAGCTGCATATTTGGTGTCTCTCAAAACCTGCCGCTACTCTCTGAGCTAT 1387
QY 1590 GGGTGTGATCCTTAACATATATATCCCTTCAGGGCTTTTACTGCTTTACAGATGGGAAATGAA 1649
Db 1388 GGGTGTGATCCTTAACATATATATCCCTTCAGGGCTTTTACTGCTTTACAGATGGGAAATGAA 1447
QY 1650 AATGTACAGCAACCTCCCAAGAGGGTATCTATTAGTAAATTCAGAGGCGACAGACAA 1709
Db 1448 AATGTACAGCAACCTCCCAAGAGGGTATCTATTAGTAAATTCAGAGGCGACAGACAA 1507
QY 1710 TTGCTGGAAGCTGCAAAAGGCTGGAGATGTGAAACTGTAAAAAACTGTGTACTGTTTCAAG 1769
Db 1508 TTGCTGGAAGCTGCAAAAGGCTGGAGATGTGAAACTGTAAAAAACTGTGTACTGTTTCAAG 1567
QY 1770 AGTCTCAACTGCAGAGACATTTGAAGGGCTGAGTCTACACCACCTTCAATTTTGGAGCTGGG 1829
Db 1568 AGTCTCAACTGCAGAGACATTTGAAGGGCTGAGTCTACACCACCTTCAATTTTGGAGCTGGG 1627
QY 1830 TATAACAGAGTGTCCGTGGTGGAAATATCTCTACAGCATGGAGCTGATGTGCATGCTAAA 1889
Db 1628 TATAACAGAGTGTCCGTGGTGGAAATATCTCTACAGCATGGAGCTGATGTGCATGCTAAA 1687
QY 1890 GATAAGGAGGCTTTGACCTTTGCAACATGCAATGTTCTTTATGGACATTTATGAAGTTGCA 1949
Db 1688 GATAAGRRGSCCTTGATCCTTTGCAACATGCAATGTTCTTTATGGACATTTATGAAGTTGCA 1747
QY 1950 GAACTTCTTCTTAAACATGAGCAGTAGTTAATGTAGTCTGATTTATGGAATTTTACACCT 2009
Db 1748 GAACTTCTTCTTAAACATGAGCAGTAGTTAATGTAGTCTGATTTATGGAATTTTACACCT 1807
QY 2010 TTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 2069
Db 1808 TTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 1867
QY 2070 GCAGACCTTACCAAAAACACAGGATGGAATACTCCTTTGGATCTTTTAAAGATGGA 2129
Db 1868 GCAGACCTTACCAAAAACACAGGATGGAATACTCCTTTGGATCTTTTAAAGATGGA 1927
QY 2130 GATACAGATATTCAGATCTGTTAGGGAGATGCGAGCTTTGCTAGATGCTGCCAAGAAG 2189
Db 1928 GATACAGATATTCAGATCTGTTAGGGAGATGCGAGCTTTGCTAGATGCTGCCAAGAAG 1987
QY 2190 GGTGTTTATGCCAGATGGAAGAGTTGTCTCTCTCTGATAATGTAATTTCCCGCATACC 2249
Db 1988 AGTGTGTTAGCCAGATGGAAGAGTTGTCTCTCTCTGATAATGTAATTTCCCGCATACC 2047
QY 2250 CAAGGCGAGACATTCACACCTTTACATTTAGCAGCTGGTTATTAATTTAGAGATTGCA 2309
Db 2048 CAAGGCGAGACATTCACACCTTTACATTTAGCAGCTGGTTATTAATTTAGAGATTGCA 2107
QY 2310 GAGTATTTGTTACACACGAGCTGATGTAATGCCCAAGCAAAAGGAGGACTTATTCTCT 2369
Db 2108 GAGTATTTGTTACACACGAGCTGATGTAATGCCCAAGCAAAAGGAGGACTTATTCTCT 2167
QY 2370 TTACATAATGTCAGCATCTTACGGGCATGATGATGTAGCAGCTCTACTTAATTAAGATTAAT 2429
Db 2168 TTACATAATGTCAGCATCTTACGGGCATGATGATGTAGCAGCTCTACTTAATTAAGATTAAT 2227
QY 2430 GCATGTCTCAATCCCGAGCAAAATGGGCTTTTACACCTTTTGCACAAAGCAGGCCAAAAG 2489
Db 2228 GCATGTCTCAATCCCGAGCAAAATGGGCTTTTACACCTTTTGCACAAAGCAGGCCAAAAG 2287
QY 2490 GGACCAACACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 2549
Db 2288 GGACCAACACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 2347
QY 2550 CAGGAAGGCAACACCTTTTAGATTAGTTTACGGGATGATGTACGCGCTCTTCTTGACA 2609
Db 2348 CAGGAAGGCAACACCTTTTAGATTAGTTTACGGGATGATGTACGCGCTCTTCTTGACA 2407
QY 2610 GCAGCCATGCCCCCATCTGCTGCTGCTGCTTGTACAAAGCCTCAAGTGTCAATGTGTG 2669
Db 2408 GCAGCCATGCCCCCATCTGCTGCTGCTGCTTGTGTACAAAGCCTCAAGTGTCAATGTGTG 2467

QY 2670 AGAAGCCAGGAGCCACTGCGAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2729
DB 2468 AGAAGCCAGGAGCCACTGCGAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2527
QY 2730 TCTGCAGCCAGCAGCTTGGAACTTATCTGGGAGTTTTCAGAACTGCTCTTCAGTAGTT 2789
DB 2528 TCTGCAGCCAGGAGCTTGGAACTTATCTGGGAGTTTTCAGAACTGCTCTTCAGTAGTT 2587
QY 2790 AGTTCAAGTGGAAACAGAGGCTTCCAGTTTGGAGAAAAAGGAGTTCCAGGAGTAGAT 2849
DB 2588 AGTTCAAGTGGAAACAGAGGCTTCCAGTTTGGAGAAAAAGGAGTTCCAGGAGTAGAT 2647
QY 2850 TTTAGCAATACTCAATTCGTAGGAATCTTGGACTTGGACCTTAATGGATATATTTCAG 2909
DB 2648 TTTAGCAATACTCAATTCGTAGGAATCTTGGACTTGGACCTTAATGGATATATTTCAG 2707
QY 2910 AGAGAACAGATCACTTTGGATGATATTAGTTGAGATGGGGCACAGGAGCTGAAGAGATT 2969
DB 2708 AGAGAACAGATCACTTTGGATGATATTAGTTGAGATGGGGCACAGGAGCTGAAGAGATT 2767
QY 2970 GGAATCAATGCTTATGGACATAGGCACAACTTAATTAAGGAGTGGAGAGCTTATCTCC 3029
DB 2768 GGAATCAATGCTTATGGACATAGGCACAACTTAATTAAGGAGTGGAGAGCTTATCTCC 2827
QY 3030 GGACACAAAGGCTTAAACCCATATTAACTTTGAACACTCTGGTAGTGGAACTTCTT 3089
DB 2828 GGACACAAAGGCTTAAACCCATATTAACTTTGAACACTCTGGTAGTGGAACTTCTT 2887
QY 3090 ATAGATCTCTCTCTCATATAAGAGTTTCACTCTGTGGAGGAAGAGATGCAAACTACA 3149
DB 2888 ATAGATCTCTCTCTCATATAAGAGTTTCACTCTGTGGAGGAAGAGATGCAAACTACA 2947
QY 3150 GTTCGAGACACAGAGATGGAGGTCATGCAAGTGGAACTTCAACAGATACAAATTTCTC 3209
DB 2948 GTTCGAGACACAGAGATGGAGGTCATGCAAGTGGAACTTCAACAGATACAAATTTCTC 3007
QY 3210 AGATTTCAAGGTTTGTATACAGAACTTATGGAAAGATACACACCGGAGAAAGAA 3269
DB 3008 AGATTTCAAGGTTTGTATACAGAACTTATGGAAAGATACACACCGGAGAAAGAA 3067
QY 3270 GTTCTTGAAGAAACACACACCATGCAATGAACGAATGCTATTTTCATGGGTCTCTTTT 3329
DB 3068 GTTCTTGAAGAAACACACACCATGCAATGAACGAATGCTATTTTCATGGGTCTCTTTT 3127
QY 3330 GTCAATGCAATTTATCCAAAGGCTTTGATGAAGGATCGCTACATAGGTGGTAGTTT 3389
DB 3128 GTCAATGCAATTTATCCAAAGGCTTTGATGAAGGATCGCTACATAGGTGGTAGTTT 3187
QY 3390 GGAGCTGGCATTTATTTTGTGCTGAAACTCTTCCAAAGCAATCAATATATATGAAAT 3449
DB 3188 GGAGCTGGCATTTATTTTGTGCTGAAACTCTTCCAAAGCAATCAATATATGAAAT 3247
QY 3450 GGAGGAGGTACTGGGTGTCAGTTTCAAAAGACAGATCTTGTACATTTGCCACAGGCAG 3509
DB 3248 GGAGGAGGTACTGGGTGTCAGTTTCAAAAGACAGATCTTGTACATTTGCCACAGGCAG 3307
QY 3510 CTGCTCTTTGGCGGTAACTTTGGAAAGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 3569
DB 3308 CTGCTCTTTGGCGGTAACTTTGGAAAGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 3367
QY 3570 GCATTTCTCTCCAGGTGTCATCACTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3629
DB 3368 GCATTTCTCTCCAGGTGTCATCACTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3427
QY 3630 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCTCAGTAGTTTAAATTAATTAC 3689
DB 3428 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCTCAGTAGTTTAAATTAATTAC 3487
QY 3690 CAGATTATAGGCTGAGGTATGGTGGATGAATAATAGTTATTTTAAAGAACTAATTC 3749
DB 3488 CAGATTATAGGCTGAGGTATGGTGGATGAATAATAGTTATTTTAAAGAACTAATTC 3547

QY 3750 CACTGAACCTAAATCATCAAGCAGCAGTGGCTCTACGTTTACTCTCTTTCTGCTGAAA 3809
DB 3548 CACTGAACCTAAATCATCAAGCAGCAGTGGCTCTACGTTTACTCTCTTTCTGCTGAAA 3607
QY 3810 AAAA 3813
DB 3608 AAAA 3611
RESULT 2
US-09-350-982C-4
; Sequence 4, Application US/09350982C
; Patent No. 6455290
; GENERAL INFORMATION:
; APPLICANT: Berthelsen, Jens
; APPLICANT: Toma, Salvatore
; APPLICANT: Isacchi, Antonella
; TITLE OF INVENTION: Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods R
; FILE REFERENCE: PHRM-0043
; CURRENT APPLICATION NUMBER: US/09/350,982C
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 3498
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: PCR Primers
; NAME/KEY: CDS
; LOCATION: (1)..(3498)
; NAME/KEY: misc_feature
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (1102)..(1102)
; OTHER INFORMATION: n is any nucleic acid
; NAME/KEY: misc_feature
; LOCATION: (2650)..(2650)
; OTHER INFORMATION: n is any nucleic acid
US-09-350-982C-4
Query Match 91.4%; Score 3488.8; DB 4; Length 3498;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3487; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 225 ATGTCGGTCCCGCTGCGCGGGGAGCGCTGCGGAGCGCTGCGGAGCGCGCGCGCGAGGCC 284
DB 1 ATGTCGGTCCCGCTGCGCGGGGAGCGCTGCGGAGCGCTGCGGAGCGCGCGCGCGAGGCC 60
QY 285 GTGGAGCGCGCGCGCGCGAGAGCTGTTCCAGGCGTCCCGAACACGGGAGCTGGAACGAGTC 344
DB 61 GTGGAGCGCGCGCGCGCGAGAGCTGTTCCAGGCGTCCCGAACACGGGAGCTGGAACGAGTC 120
QY 345 AAGAGCTGGTGACCGCTGGAAGGTGAACAGCCCGACAGCGCGCGCGAGAAATCCACC 404
DB 121 AAGAGCTGGTGACCGCTGGAAGGTGAACAGCCCGACAGCGCGCGCGAGAAATCCACC 180
QY 405 CGCTGCACTTCGCCCGAGGTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAAT 464
DB 181 CGCTGCACTTCGCCCGAGGTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAAT 240
QY 465 GGTGCAATATGTCACAGCAGTGAATGATGGGGCGCTTATTCCTCTTCAATAATGCAATGCTCT 524
DB 241 GGTGCAATATGTCACAGCAGTGAATGATGGGGCGCTTATTCCTCTTCAATAATGCAATGCTCT 300
QY 525 TTTGGTCACTGTAAGTAGTCAATCTCTTTTCCACATGTTGCGAGACCCCAATGCTCGA 584
DB 301 TTTGGTCACTGTAAGTAGTCAATCTCTTTTCCACATGTTGCGAGACCCCAATGCTCGA 360
QY 585 GATAATTTGAATTTATCTCTCTCCATGAAGCTGCAATTTAAAGAAAGATTTGATTTTTC 644
DB 361 GATAATTTGAATTTATCTCTCTCTCCATGAAGCTGCAATTTAAAGAAAGATTTGATTTTTC 420

Db 2581 GAGGGTCTCCAGTTTGGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2640
Qy 2865 TTCGTAAGGAATCTTGGACTTCAGCACCTAATGGATATATTTGAGAGAGACAGATCACT 2924
Db 2641 TTCGTAAGGNATCTTGGACTTCAGCACCTAATGGATATATTTGAGAGAGACAGATCACT 2700
Qy 2925 TTGGATGATATTAGTTGAGATGGGGCAGAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2984
Db 2701 TTGGATGATATTAGTTGAGATGGGGCAGAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2760
Qy 2985 GGACATAGGCACAACTAATTAAGAGAGTCGAGAGACTTATCTCCGGACACAAAGGCTTT 3044
Db 2761 GGACATAGGCACAACTAATTAAGAGAGTCGAGAGACTTATCTCCGGACACAAAGGCTTT 2820
Qy 3045 AACCCATATTAACTTTGAACACCTCTGCTAGTGAACAAATCTTATAGATCTGCTCCT 3104
Db 2821 AACCCATATTAACTTTGAACACCTCTGCTAGTGAACAAATCTTATAGATCTGCTCCT 2880
Qy 3105 GATGATAAAGAGTTTCAGTCTCTGGAGGAAGAGATCCAAAGTACAGTTCCGAGAGCACAGA 3164
Db 2881 GATGATAAAGAGTTTCAGTCTCTGGAGGAAGAGATGCAAAAGTACAGTTCCGAGAGCACAGA 2940
Qy 3165 GATGGAGTCATGCGAGGTGGAATCTTCAACAGATACATATTTCTCAAGATTCAGAAGTTT 3224
Db 2941 GATGGAGTCATGCGAGGTGGAATCTTCAACAGATACATATTTCTCAAGATTCAGAAGTTT 3000
Qy 3225 TGTAACAAGAACTATGGGAAGATACACTCCAGGAGAAAGAGTTCTCGAAGAAAC 3284
Db 3001 TGTAACAAGAACTATGGGAAGATACACTCCAGGAGAAAGAGTTCTCGAAGAAAC 3060
Qy 3285 CACAACCATGCCAATGAACGAATGCTATTTTCATGGGTCTCCTTTTGTGAATCAATATTC 3344
Db 3061 CACAACCATGCCAATGAACGAATGCTATTTTCATGGGTCTCCTTTTGTGAATCAATATTC 3120
Qy 3345 CACAAGGCTTTGATCAAAAGGATGCGGTACATAGTGGTATGTTTGGAGCTGGCATTTAT 3404
Db 3121 CACAAGGCTTTGATCAAAAGGATGCGGTACATAGTGGTATGTTTGGAGCTGGCATTTAT 3180
Qy 3405 TTTGCTGAAACCTCTTCCAAAAGCAATCAATATGTATATGGAATTCGAGAGGCTACTGG 3464
Db 3181 TTTGCTGAAACCTCTTCCAAAAGCAATCAATATGTATATGGAATTCGAGAGGCTACTGG 3240
Qy 3465 TGTCAGTTTCACAAAGACAGATCTTGTATACATTTGCCACAGGAGCTGCTTTTGGCGG 3524
Db 3241 TGTCAGTTTCACAAAGACAGATCTTGTATACATTTGCCACAGGAGCTGCTTTTGGCGG 3300
Qy 3525 GTAACTTGGGAAGTCTTCTGCAATTCAGTGCATGCAATGAAATGGCAGATCTCTCCA 3584
Db 3301 GTAACTTGGGAAGTCTTCTGCAATTCAGTGCATGCAATGAAATGGCAGATCTCTCCA 3360
Qy 3585 GGTCACTCACTCACTGCTAGGCCCCAGTGAATGGCTTAGCATTTAGCTCAATATGTT 3644
Db 3361 GGTCACTCACTCACTGCTAGGCCCCAGTGAATGGCTTAGCATTTAGCTCAATATGTT 3420
Qy 3645 ATTTACAGAGGAGAACAGGCTTATCTGAGTATTTAAATTTACTTACAGATTAAGGCTT 3704
Db 3421 ATTTACAGAGGAGAACAGGCTTATCTGAGTATTTAAATTTACTTACAGATTAAGGCTT 3480
Qy 3705 GAAGGTATGGTCATGGA 3722
Db 3481 GAAGGTATGGTCATGGA 3498

RESULT 3
US-09-196-387-1
; Sequence 1, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRP1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/196,387
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/095,225
FILING DATE: June 10, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4134 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-09-196-387-1

Query Match 49.4%; Score 1885; DB 4; Length 4134;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 2496; Conservative 0; Mismatches 940; Indels 27; Gaps 2;
Qy 252 GGAGGGGCTCGCGAGCGCCGCGGAGCGCTGAGCGCGCGCGCGAGAGCTGTTTC 311
Db 507 GGGGAGGAGGAGCTGGGAGAGGGTCCAGCAGTGAAGGGGCGGCTACGGGAACTGCTG 566
Qy 312 GAGGGTCCGCAAGGGGAGCTGGAAGAGTCAAGAGGCTGGTGACGCTGAGAGGTTG 371
Db 567 GAGGGCTGCGCAATGGGAGCTGTCCCGGTAAGAGGCTGGTGAGCGCGGCAACGTA 626
Qy 372 AACAGCCGCGACAGCGCGGCGGAGGAAATCCACCGCTGACATTCGCGCGAGGTTTGGG 431
Db 627 AATGCAAGGAGATGTCGCGCGCGGAGGCTTCTCCCTGCACTTCGCTGCAAGTTTGA 686
Qy 432 CGGAAGAGCTAGTTGAATATTTGCTCAGAAATGGTGAATGTCGAAGCAGCTGATGAT 491
Db 687 AGGAAGGATGTTGAGAACACTTACATGATGGTGTGCTAATGTCAGGCTGCGTATGAT 746
Qy 492 GGGGGCTTATTCTCTTCATATAATGCTGCTTTTGGTTCATGCTGAAGTAGTCAATCTC 551
Db 747 GGAGTCTCATCCCGCTTCATATAAGCCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806
Qy 552 CTTTTCGACATGCTGACAGCCCCCAATGCTCGAGATAATTTGAATTAATACCTCTCTCCAT 611
Db 807 TTATTGTCCCAAGGAGCTGATCCAAATCCAGGGGATAACTGGAATATACACCTCTCTGAT 866
Qy 612 GAAGTCAATTAAGGAGAAAGATTCATGTTTGCATTTGCTGTTTACAGCATGGAGCTGAG 671
Db 867 GAAGTCAATTAAGGAGAAAGATTCATGTTTGCATTTGCTGTTTACAGCATGGAGCTGAG 926
Qy 672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTTAGCAGATCCATCTGCCAAA 731
Db 927 CCAACCATCCGAAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCTCTCAGCAAAA 986

Db	3126	GAACACCTTTCCGGGATATCTTTGAACAGACAGAGATTACACTAGATGTTGGCTGATATG	3185
Qy	2946	GGSCACAAGGCTGAAGGAGATTGGAATCAATGCTTTATGGACATAGGCACAACATAAT	3005
Db	3186	GGTCATCAGAGTTGAAAGAAATAGGCATCAATGCATATGGCACGCCACAATTAATC	3245
Qy	3006	AAAGGAGTCGAGAGACTTATCTCCGACACAAGGTCTTAACCCATATTTAACTTTGAAC	3065
Db	3246	AAAGGAGTAGAAGACTCTTAGTGGACAACAAGGCACCAATCCTTTATTTGACTTTTCAC	3305
Qy	3066	ACCTCTGGTAGTGGAAACAAATCTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTTCACTCT	3125
Db	3306	TGTGTTTAATCAGGGAACGATTTTGTCTGGATCTTCTCCAGAAGATAAAGAAATATCAGTCA	3365
Qy	3126	GTCCGGAAGAGATGCAAAATACACTTCGAGAGCACAGATGAGGTCTATGCAGGTGGA	3185
Db	3366	GTGGAAGAAGATGCAAAAGTACTATTTCGAGACACAGAGATGGTATGCTGGCGC	3425
Qy	3186	ATCTTCAACAGATACAAATATCTCAAGATTTCAGAAGTTTGTAAACAAGAACTATGGAA	3245
Db	3426	ATCTTCAACAGATACAAATGCTATTCGAATTCAAAAGTTGTCAACAAGAAGTTGAGGAG	3485
Qy	3246	AGATACACTCACCGGAGAAAGAACTTTCTGAAGAAAACCAACCAATGCCAATGAACGA	3305
Db	3486	CGGTTCTGCCACGACAGAAGGAAGTGTCTGAGGAGAATCAACACCATCACAATGAGCGC	3545
Qy	3306	ATGCTATTTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAGGCTTTGATGAAGG	3365
Db	3546	ATGTTGTTTTCATGGTTCTCCTTTCAATTATGCCATTATCAATAAAGGTTTGATGAGCGA	3605
Qy	3366	CATGCGTACATAGTGGTATGTTTCGAGCTGCATTTATTTTGTCTGAAAACCTCTTCCAAA	3425
Db	3606	CATGCATACATAGGAGGAATGTTTGGGGCGGGATTTATTTTGTGAAAACCTCTCAAAA	3665
Qy	3426	AGCAATCAATATGATATATGGAATTGGAGGAGTACTGGGTGCCAGTTCACAAAGACAGA	3485
Db	3666	AGCAACCAATATGTTTATGGAATTGGAGGAGAACAGGCTGCCCTACACACAAGGACAGG	3725
Qy	3486	TCCTGTGTACATTTGCCACAGGAGCTGCTCTTTGCCGGGTAAACCTTTGGGAAAGCTTTTC	3545
Db	3726	TCATGCTATATATGTCACAGACAAATGCTCTCTCTAGAGTGACCCCTTGGGAAATCCCTT	3785
Qy	3546	CTGCAGTTCACTCAATGAAATGGCACATTTCTCCTCCAGGTCACTCACTCAGTCAGTGGT	3605
Db	3786	CTGCAGTTTAGCAACCAATGAAATGGCCCAACGCCCTCCAGGGCACCTCAGTCATTTGGT	3845
Qy	3606	AGGCCAGTGTAAATGGCCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAACAGGCT	3665
Db	3846	AGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAACAGGCA	3905
Qy	3666	TATCCTCAGTATTTAATTTACTTACACAGATTTATGAGGCCTGAAG	3708
Db	3906	TACCGAGATCTTATCTACTTTACCAAGATCATGAAGCCACAAG	3948

RESULT AND DISCUSSION

```

RESOLUTION 4
US-09-196-387-7
; Sequence 7, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

```


Db	3207	AGAGCTTTGCTGATAGATGCCATGCCCCAGAGGCCTTACCTACCTGTGTTTTAAACCTCAG	3266
Qy	2655	GTGCTCAATGGTGTGAGAAGCCCGAGGAGCCACTGACGATGCTCTCTCTTCACGGTCCATCT	2714
Db	3267	G-----CTACTGTAGTGAGTGCCTCTCTGATCTCACCGACATCC	3305
Qy	2715	AGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTCACAACTTATCTGGAGTTTTCAGAA	2774
Db	3306	ACCCCTCTGCTCTCGGCTGCCAGCAGATAGACAACCTCACTGGCCCTTTACGACAG	3365
Qy	2775	CTGCTTTCAGTAGTTAGTTCAAGTGGAAACAGAGGTCCTTCCAGTTTGGAGAAAAAG---	2831
Db	3366	TTGGCCGTAGGAGGCCCTCCAATGCAGGGGATGSCGCCGGGAACAGAAAGGAAGAA	3425
Qy	2832	---GAGGTTCCAGGAGTAGATTTTAGCTAATACTCAATTCGTAGGAACTCTTGACATTGAG	2888
Db	3426	GGAGAAGTTGCTGGCTGTGACATGAATATGACGCCAATTTCTAAAGAGCCTTGCCCTTGA	3485
Qy	2889	CACCTAATGGATATTTTGGAGACAACAGATCACTTTGGATGTATTAGTTGAGATGGGG	2944
Db	3486	CACCTTCGGGATCTTTGAAACACACAGATTACACTAGATGTGTGGCTGATATGGCT	3545
Qy	2949	CACAAGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACTAATATAA	3008
Db	3546	CATGAAGAGTTGAAGAAAAATAGGCATCAATGCATATGGCACGCCCAAAATTAATCAA	3605
Qy	3009	GGAGTCAGAGACTTATCTCCGAGACAACAGGTCCTTAACCCATATTTAACTTTGAACACC	3068
Db	3606	GGAGTAGAAAGACTCTTAGTGGACACAAAGGCCCAATCCTTATTTCACCTTTCACTGT	3665
Qy	3069	TCTGGTAGTGGAACTTCTTATAGATCTGTCTCCTGATGAATAAGAGTTTCAGTCTGTG	3128
Db	3666	GTTAATCAGGGAACGATTTTGTGTGATCTTGCTCCAGAAAGATAAAGAAATATCACTG	3725
Qy	3129	CAGGAACAGATGCAAAAGTACAGTTTCAGAGCACAGAGATGCAGCTCATGCAGGTGAAATC	3188
Db	3726	GAAGAAGAGATGCAAAAGTACTATTTCGGAACACAGAGATGGTGTATGCTGGCGGATC	3785
Qy	3189	TTCAACAGATACAAATATCTCAAGATTTCAGAAGTTTGTAAACAAGAACTATCGGAAAGA	3248
Db	3786	TTCAACAGATACAAATGTCAATTCGAATTCAAAAGTTGTCAACAAGAAGTTGAGGAGCGG	3845
Qy	3249	TACACTCACGGGAGAAAGAAGTTTCTGAAGAAAACACACCAATGCCATGACCAATG	3308
Db	3846	TTCTGCCACCGACAGAGGAAGTGTCTGAGGAGAATCAACACCTACAAATGAGCGCATG	3905
Qy	3309	CTATTTTCATGGGTCTCCTTTTGTGAATGCAATATTCACAAAGGCTTTGATGAAAGCAT	3368
Db	3906	TGTTTTCAATGGTCTCCTTTTCATTAATGCCATTATTCATAAAGGTTTGATGACGCAT	3965
Qy	3369	CGGTACATAGTGGTATGTTTGGAGTGCATTTTATTTTGCTGAAACTCTTCCAAAAGC	3428
Db	3966	GCATACATAGAGGAATGTTTGGGCGCGGATTTATTTTGTGTAACCTCCCTCAAAAAGC	4025
Qy	3429	RATCAATATGTTATGAAATGGAGAGTACTGGGTGCCAGTTTCAGTTTCACAAAGACAGATCT	3488
Db	4026	RACCAATATGTTTATGAAATGGAGAGAACAGGCTGCCCTACACACAGGACAGGTCA	4085
Qy	3489	TGTTACATTTGCCACAGGCAGCTCTCTTTTGCGGGTAACTTGGGAAAGCTTTTCTGTG	3548
Db	4086	TGCTATATATGTCACACACAAATGCTCTCTGTAGAGTGACCTTGGGAAATCCTTTCTG	4145
Qy	3549	CAGTTTCAGTCAATGAAAATGGCACATTTCTCTCCAGTCTCATCTCACTCACTGCTGAGG	3608
Db	4146	CAGTTTAGCCCATGAAAATGGCCAGCGGCTCCAGGGCAGCACTCACTGATTTGGTAGA	4205
Qy	3609	CCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAACACAGGCTTAT	3668
Db	4206	CCGACCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGACACAGGCATC	4265
Qy	3669	CCTGAGTATTTAATTACTTACCAGATTATGAGGCCCTGAAG	3708
Db	4266	CCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG	4305

QY 2399 AGATGTAGCAGCTCTACTAATAAAGTATATAATGCATGTGTCAATGCCAGGCAAAATGGGC 2458
DB 2967 TGACATAGCGGCTTTATTCATATAATACACACAGCTGTGTAAATGCAACAGATAAGTGGC 3026
QY 2459 TTTCACACCTTTTGCAGGAGAGCCCAAAAGGAGGACACAGCTTTGTGCTTTGCTTGT 2518
DB 3027 GTTTACTCCCTCCATGAAGCAGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3086
QY 2519 AGCCCATGAGCTGACCCGACCTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTAGT 2578
DB 3087 AGCCCATGAGCTGACCCGACCTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTAGT 3146
QY 2579 TTTCAGGAGTGTGACGCGCTCTTCT 2605
DB 3147 AACAGCTGACGATATCAGAGCTTTGACATGTTGACATAGCGGCTTTATTCATAAATAACAA 3206
QY 2606 2605
DB 3207 CAGCTGTAAATGCAACAGATAAGTGGCGTTTACTCCCTCCATGAAGCAGCCAGAA 3266
QY 2606 -GACAGCAGCCTATGCCCATCTGCTCTGCCCTCTGTTTACAAGCCTCAAG- 2655
DB 3267 AGGAAGCAGCAGCTGTGCGCCTCTCTAGCGATGTGAGACCCACCATGAAGAA 3326
QY 2656 2662
DB 3327 CCAGGAAGCCACAGCCTCTGATCTGGCATGCGCAACAGCTGACGATATCAGAGCTTTGCTGAT 3386
QY 2663 TGTGTGAGAGCCCGAGGAGCCACTGC- 2689
DB 3387 AGATGCGATGCCCGCAGAGGCGCTTACCTACTGTTTAAACCTCAGGCTACTGTAGTAG 3446
QY 2690 AGATGCTCTCTCTCAGGTCTTACAGCTTACAGCCATCAAGCCTTTCTGAGCCAGCAGCTTGA 2749
DB 3447 TGCCCTCTGATCTACACAGCATCCACCCCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 3506
QY 2750 CAATCTATCTGGAGGTTTTCAGAACTGTCTTCAAGTGTAGTTAGTTCAAGTGGAAACAGAGG 2809
DB 3507 CAACCTCAGTGGCCCTTAGCAGAGTTGGCCGTAGGAGGAGCTTCAATGCGAGGATGG 3566
QY 2810 TGCTTCAGCTTTGGAGAAAGAG- - - - -GTTCCAGAGGTAGATTTTAGCATAACTCA 2863
DB 3567 CGCCGGGGGAACAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3626
QY 2864 ATTCTGAAGCAATCTTGGAGCTTGCAGCCTTAATGGATATATTGAGAGAGACAGATCAC 2923
DB 3627 ATTCTGAAGCAATCTTGGAGCTTGCAGCCTTAATGGATATATTGAGAGAGACAGATCAC 3686
QY 2924 TTTGGATGTATTAGTTGAGATGGGACAGAGGAGCTGAAGGAGATTGGAATCAATGCTTA 2983
DB 3687 ACTAGATGTGTTGCTGATGCTGATGAGAGGTTGAAGAAATAGGATCAATGCTTA 3746
QY 2984 TGGACATAGGCAAACTAATTAAGAGGTGAGAGCTTATCTCCGGACACAAAGTCT 3043
DB 3747 TGGGACCGGCAAACTAATTAAGAGGTGAGAGCTTATAGGTTGGACAAACAGGAC 3806
QY 3044 TAACCCATATTAACTTTGACACCTCTGCTAGTGAACAAATCTTATAGATCTGCTCC 3103
DB 3807 CAATCTTATTTGACCTTTTACCTGCTGTTAATCAGGAAAGGATTTTCTGCTGCTGCTCC 3866
QY 3104 TGATGATAAAGAGTTTCACTCTGAGGAGAGATGCAAGGATTTGAGAGGACACAG 3163
DB 3867 AGAGATAAAGATATACAGTCTAGTGGAGAGAGATGCAAGTACTATTTCGAGAACACAG 3926
QY 3164 AGATGAGGTGATGACAGGTGGAATCTTCAACAGATCAATATTCTCAAGATTGAGAGGT 3223
DB 3927 AGATGAGGTGATGCTGGGCGCTCTTCAACAGATCAATATTCTCAAGATTGAGAGGT 3986
QY 3224 TTGTAAACAAGAACTATGGGAAGATACACTCACCAGGAGAAAGAGTTTCTGCAAGAAA 3283
DB 3987 TGTCACAAGAAAGTTGAGGAGGCTTCTGCCACCGAGAGGAGGTGCTGAGAGAA 4046

QY 3284 CCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTAT 3343
DB 4047 TCACAACCATCACAAATGAGCGCATGTTGTTTCATGGTCTCCTTTTCAATTAATGCAATTAT 4106
QY 3344 CCACAAGGCTTTTCATGAAGGCGATGCTACATAGGTGGTATGTTTGGAGCTGGCATTTA 3403
DB 4107 TCATAAAGGTTTTCATGAGCGCATGCTACATAGGAGGAATGTTTGGGCGGGAATTA 4166
QY 3404 TTTTGTCTGAAACCTTCTCCAAAGCAATCAATATGATATGGAATTTGGAGAGGAGTACTGG 3463
DB 4167 TTTTGTCTGAAACCTTCTCCAAAGCAATCAATATGATATGGAATTTGGAGAGGAGTACTGG 4226
QY 3464 GTGTCCAGTTTCCAAAGCAATCAATATGATATGGAATTTGGAGAGGAGTACTGG 3523
DB 4227 CTGCGCTACACAGGAGGAGTCTATATATATGATATGGAATTTGGAGAGGAGTACTGG 4286
QY 3524 GGTAACTTGGGAAGTCTTCTGCGAGTTTCAGTGCAGTGAATGAATGCGACATTTCTCCTCC 3583
DB 4287 AGTGAACCTTGGGAATCTTCTGCGAGTTTCAGTGCAGTGAATGAATGCGCGCTCC 4346
QY 3584 AGGTCACTCACTCAGTCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3643
DB 4347 AGGTCACTCACTCAGTCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4406
QY 3644 TATTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3703
DB 4407 CATCTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4466
QY 3704 TGAAG 3708
DB 4467 AGAAG 4471
RESULT 6
US-09-196-387-12
; Sequence 12, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196.387
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/095,225
; FILING DATE: June 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 base pairs

QY 2402 TGTAGCAGCTCTACTAATAAAGTATATATGATGTGTCAATGCCACGACAAATGGGCTTT 2461
Db 1182 AATTGTGGAAGTGTACTGAAATACCAAGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1241
QY 2462 CACACCTTTGACCAAGAGCAGCCCAAGGAGGACGACACAGCTTTGTCTTTGTGTAGC 2521
Db 1242 CTCACCGCTGCATGCTGCTTTTATGGTGTATTAACATTTGTCATCTATTACTACA 1301
QY 2522 CCATGGAGCTCACCGGACTCTTTAAATATCAGGAAGGACAAACACCTTTAGATTTTTC 2581
Db 1302 ACAAGGTCTAATGAGATGTGGTACAGTACGGGTGCAAGCCCTCTTCATTTAGCTGC 1361
QY 2582 A 2582
Db 1362 A 1362

RESULT 8
US-09-031-485-36/c

; Sequence 36 Application US/09031485

; Patent No. 5824306

; GENERAL INFORMATION:

; APPLICANT: Tang, Liang

; APPLICANT: Blehm, E. Scot

; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND

; TITLE OF INVENTION: USES THEREOF

; NUMBER OF SEQUENCES: 85

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hecla Corporation

; CITY: Fort Collins

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80525

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Windows 95

; SOFTWARE: WordPerfect for Windows, Version 7.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/031.485

; FILING DATE:

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/847.429

; FILING DATE: 24-APR-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Verser, Carol Talkington

; REGISTRATION NUMBER: 37,459

; REFERENCE/DOCKET INFORMATION: HW-5

; TELEPHONE: 970/493-7272

; TELEFAX: 970/484-9505

; INFORMATION FOR SEQ ID NO: 36:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5235 nucleotides

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; US-09-031-485-36

Query Match 2.4%; Score 93; DB 1; Length 5235;

Best Local Similarity 45.0%; Pred. No. 6.6e-14;

Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

QY 1392 GGAGCAACATCAATGAAGACTAAAGATTTCTTGACTCTCTCGACGTGGCATCTGAG 1451

Db 5058 GGCACCGATATCAACACATGCAATGCAATGCGCTTAATGCAATTCATCTGCGCTCCAAA 4999

QY 1452 AAAGCTCATAAATGATGTTGTTGAAGTAGTGGTCAAAACATGAAGCAAGGTTAATGCTCTG 1511
Db 4998 GAAGGTCAATCATGAAGTGGTCCCGAACTTCTGAAAAGAAAGCAGATCTTGATGCTGCC 4939
QY 1512 GATAATCTTTGGTCAGACTTCTCTACACAGAGTGCATATTTGGTGCATCTCAAAACCTGC 1571
Db 4938 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTTGGCAGGACAAGAACTAATCTGC 4879
QY 1572 CGCTACTCTCTGAGCTATGGGTGATCTTAACATTATATCCCTTCAGGGCTTTTACTGCT 1631
Db 4878 ACAGTACTTGTTCGAAATGGTCTAATTTAGCTACAATCACTAAACGGTTTTACACCA 4819
QY 1632 TTACAGATGGGAAATGAAATGTACAGCAACTCTCCAAGAGGGTATCTCAATAGGTAAT 1691
Db 4818 CTTTACATGGCTGCACAAGAAATCACGAATCTGTTGTAGCTATCTTCTTGCACCAAT 4759
QY 1692 TCAGAGGCACAGACAAATTTGCTGGAGCTGCAAGGCTGGAGATGT-----CGAAAC 1744
Db 4758 GCCAATCAAGCTTTAAGTACAGAAAGACGGTTTTTACGCCACTGGCAGTTGCCCTTGAACAA 4699
QY 1745 TGTAAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1804
Db 4698 GGTACAGATCGTGTGGTCTGTTTGTGTTGAAATGCACCGCGGAAAGTG-CGCTT 4640
QY 1805 TACACCACTTCATTTTTCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGTCTACA 1864
Db 4639 GCCAGCACTGCATATTTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4580
QY 1865 GCATGGAGTGTGTCATGCTAAAGATAAAGAGGCGTTGTACCTTTGCACAAATGCATG 1924
Db 4579 AAATGAGCATAACTCGGATGTCGTTGAAAAGCGCTTTACTCCGCTTCATATCCGCCG 4520
QY 1925 TTCTTATGGACATTATCAAGTTCGACAACTTCTTGTAAACATGGAGCAGTAGTTAATGT 1984
Db 4519 TCATTTGAAATGAGAACGTCAGCACTGCTACTCGAAGAGGAGCCAAATGTGAATTA 4460
QY 1985 AGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAAT 2044
Db 4459 CCAAGCGAGACATAAACAATAGTCCGTTACACGTTGCAACAAAATGGGTCGTACAAACAT 4400
QY 2045 TTGCAAACTTCTGTCAGCATGGTGCAGACCCCTACCACAAAAGAGGATGGAATAC 2104
Db 4399 GGTTCGTTTATGTTGGCTCATGGGCGCTAATTTGACTGTCGCACACGTAATTTACTAAC 4340
QY 2105 TCCTTTGGATCTTGTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
Db 4339 ACATTTACACTGCTCTCTGTCAGTCATGATCAAGTTGTTGATTTGTTGCTGAAA 4280
QY 2162 TGCACTTTGCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGAGTGAAGAGTTGCTTC 2221
Db 4279 AGAGCTCCAATCAGTCTAAGACAAAAAATGGTTTGGCTCCCTTTACATATGCGACACA 4220
QY 2222 TCCTGATAATGTAATTTGCCCGCATACCCAAAGGACAGACATTCACACCTTTTACATTTAGC 2281
Db 4219 GGTGGATGATGTTACTGTTGACTATCTC-----ACTCCTCTTCTATGCTGGC 4175
QY 2282 AGCTGGTTATAATAATTTAGAAATTTGACAGATATTTGTTTACAACACAGAGCTGATGTAA 2341
Db 4174 TGCTCATTTGGGACATGTCCTGTCGCTAACTTTTCTGATCGTAATGCTGACCCGAA 4115
QY 2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTTACATAATGTCAGCATCTTACGGGATGTAGA 2401
Db 4114 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGAACAAAAAATCGCATTA 4055
QY 2402 TGTAGCAGCTCTACTAATAAAGTATATGATGTGTCAATGCCACGACAAATGGGCTTT 2461
Db 4054 AATTGTCGAACTGCTACTGAAATACCAGCTGCAATCGAAGCAACTACTGAATCCGGTCT 3995
QY 2462 CACACCTTTGACCAAGCAGCCCAAGGAGGACGACAGCTTTGTGCTTTGTTGCTAGC 2521
Db 3994 CTCACCGCTGCATGCTCGTCTTTTATGGGTGCTATAAACAATGCTCATCTATTACTACA 3935
QY 2522 CCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTTAGATTTAGTTTC 2581

Db 3934 ACAAGTGCTAATGACAGATGTGGCTACAGTACCGGTGAAGCCCTCTTCAATTAGCTGC 3875
 QY 2582 A 2582
 Db 3874 A 3874

RESULT 9

US-08-847-429A-35

; Sequence 35, Application US/08847429A

; Patent No. 5827692

; GENERAL INFORMATION:

; APPLICANT: Tang, Liang

; APPLICANT: Biehm, E. Scot

; TITLE OF INVENTION: DIOPHILARIA AND BRUGIA ANKYRIN

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND

; TITLE OF INVENTION: USES THEREOF

; NUMBER OF SEQUENCES: 85

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carol Talkington Verser, Ph.D.

; ADDRESSEE: Heska Corporation

; STREET: 1825 Sharp Point Drive

; CITY: Fort Collins

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80525

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Windows 95

; SOFTWARE: WordPerfect for Windows, Version 7.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/847,429A

; FILING DATE: 24-APR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Verser, Carol Talkington

; REGISTRATION NUMBER: 37,459

; REFERENCE/DOCKET NUMBER: HW-5

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 970/493-7272

; TELEFAX: 970/484-9505

; INFORMATION FOR SEQ ID NO: 35:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5235 nucleotides

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cdna

US-08-847-429A-35

Query Match 2.4%; Score 93; DB 1; Length 5235;
 Best Local Similarity 45.0%; Pred. No. 6,6e-14;
 Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

QY 1392 GGAGCAACATCAATGAAAGACTAAAGATTCCTGACTCTCTGACGCTGCGATCTGAG 1451
 Db 178 GGCACCGATATCAACACATGCAATGCGAATGCGCTTAATGCAATTCGATCTGCGCCCTCCAAA 237

QY 1452 AAAGCTCAATAGTGTGTTGAAGTCTGTTGAACATGGAACATGAAGCAAGGTTAATGCTCTG 1511
 Db 238 GAAGTCAATAGTGTGTTGAAGTCTGTTGAACATGGAACATGAAGCAAGGTTAATGCTCTG 1511

QY 1512 GATAATCTTGGTCAGACTCTCTACACAGAGCTGCATATTTGGTGCATCTACAAACCTGC 1571
 Db 298 ACTAGAAAGGTTAACACAGCGTTACATATAGCATCTTGGCAGGACAGAACTAATCTGC 357

QY 1572 GCGTACTCTGAGCTATGGGTGATGCTTACATATATATCCCTTCAGGGTTTACTGCT 1631
 Db 358 ACAGTACTTGTGAAATGGTGTCTAATTTAAGCTACATCACTAAACGGTTTACACCA 417

QY 1632 TTACAGATGGGAAATGAAATGTACAGCACTCTCCAAAGAGGGTATCTCATTAGCTAAT 1691

Db 418 CTTTACATGCTGCACAAGAAATACGAATCTGTTGTACGCTATCTTCTTGCCCAAT 477
 QY 1692 TCAGAGGCGACAGACAATGTGCGAAGCTGCAAGGCTGGAGATGT-----CGAAAC 1744
 Db 478 GCAATCAAGCTTTAAGTACAGAAAGCGGTTTACGCCACTGGCAGTGGCTTGCCTTGAACAA 537
 QY 1745 TGTAAAAAAGCTGTGTACTCTTCAGAGTGTCAACTGCAGAGACATTCGAAGGGCTGAGTC 1804
 Db 538 GGTACAGATCGTGTGTGCTGCTGTTTGTGTTGAAATGACACGCGCGGAAAGTG-CGCTT 536
 QY 1805 TACACCACTTCAATTTTTCAGCTGGGTATAACAGAGTGTGCGTGTGGAATATCTGTACA 1864
 Db 597 GCGAGCACTGCATATTTGCTCTAANAAGATGATACAAAGCAGCTACGCTATTACTTCA 656
 QY 1865 GCATGAGAGCTGATGTCATCTGCTTAAGATAAAGGAGCGCTTGTACCTTTGCACAATCATG 1924
 Db 657 AATGAGCATAACTCGGATGTGACTTCGAAAGCGGCTTTACTCCGCTTCATATCGCCGC 716
 QY 1925 TTTCTATGGACATATGAAAGTTGCAGAACTTCTGTTAAACATGAGAGCTAGTTAATGT 1984
 Db 717 TCATATGGAATGGAACGTAGCACTGCTACTCGAAAGGGAGCAATGTGAATTA 776
 QY 1985 AGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGAAAAATATGAAAT 2044
 Db 777 CCAAGCGAGACATAACATAAGTCCGTTTACAGTTGCAACAAATGGGCTCGTACAACAT 836
 QY 2045 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAAGAGGATGGAATAAC 2104
 Db 837 GGTTCGTTATTGTTGGCTCATGGGCGCTTAATGTACTGCGCACACGTTGTTACTAAC 896
 QY 2105 TCCTTTGGATCTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGTTAGGGGAGA 2161
 Db 897 ACATTACACTGTGCTTCTGTTGAGCTCATGATCAAGTTGTTGATTTGTTGCTTCAAAA 956
 QY 2162 TGCAGCTTTGTAGATGCTGCCAAGAGGTTGTTTAGCAGAGTGAAGAAGTTGTCTTC 2221
 Db 957 AGGAGCTCCAATCAGTGTCTAAGACAAAAAATGTTTGGCTTACATATGGCAGCACA 1016
 QY 2222 TCCTGATATGTAATTTGCCGCGATACCCAGCAGACATTCACACCTTTACATTTAGC 2281
 Db 1017 GGTGATGATGTTTACTGTTGACTATCTC-----ACTCCTCTTCATGTGCG 1061
 QY 2282 AGCTGTTTATAATAATTTAGAAGTTGCAGAGTATTGTTTACAAACGAGCTGATGTGAA 2341
 Db 1062 TGCATATTGCGGACATGCTCGTTCGCTAAACTTTTGTGGATCGTAATGCTGACCCGAA 1121
 QY 2342 TGCCCAAGACAAAAGGAGGACTTATTCCTTTACATAATGCAAGCATCTTACGGGCGATAGA 2401
 Db 1122 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAATCGCATTA 1181
 QY 2402 TGTCAGCTCTTACTTAATAAGTATAAATGATGTCATGTCATGCCAGGCAAAATGGCTTT 2461
 Db 1182 AATTGTCGAACGTGCTACTGAAATACCAAGCTGCAATCGAAGCAACTACTGAATCGGTCT 1241
 QY 2462 CACACCTTTGACGAGCAGGCGGCAAGGCAACACAGCTTTGTGCTTTCTTGTGTAGC 2521
 Db 1242 CTCACCGCTGATGTCGCTGCTTTTATGGTCTATAAATGTCATCTATTACTTACA 1301
 QY 2522 CCATGGAGCTGACCGACTCTTAAAAAATCAGGAAGCAACACCTTTAGATTTTGTTC 2581
 Db 1302 ACAAGGTCTAATGACAGATGTTGGCTACAGTACGCGGTGAAAGCGCTCTTCTTACTGTC 1361
 QY 2582 A 2582
 Db 1362 A 1362

RESULT 10

US-08-847-429A-36/c

; Sequence 36, Application US/08847429A

; Patent No. 5827692

; GENERAL INFORMATION:

APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scott
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/47,429A
FILING DATE: 24-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 5235 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-847-429A-36

Query Match 2.4%; Score 93; DB 1; Length 5235;

Best Local Similarity 45.0%; Pred. No. 6.6e-14;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

QY 1392 GGAGCAACATCAATGAAAGACTAAAGAAATCTTGACCTCTGACAGTGGCATCTGAG 1451
DB 5058 GGCACCATATACACATGCAATGCAATGGCTTAATGCTATGCTATGGCTCCAA 4999
QY 1452 AAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGTTAATGCTG 1511
DB 4998 GAAGGTCATCATGAAGTGGTCCGGAATCTTGAAAGAAAGACAGATGTTGATGCTGC 4939
QY 1512 CATATCTTGGTCAGACTTCTTACACAGAGCTGCATATTTGGTCTATACAAACCTGC 1571
DB 4938 ACTAGAAGGGTAAACAGAGCTTACATATAGCATATGTCAGGACAGAAAGAAATCGTC 4879
QY 1572 CGCTACTCTGAGCTATGGTGTGATCCCAACATATATCCCTTCAGGCTTTACTGCT 1631
DB 4878 ACATGACTTGTGAAATGTTGCTTAATGTTAAGTACATCACTAAGGTTTACACCA 4819
QY 1632 TTACAGATGGGAAATGAAATGTAAGCAACTCTCTCAAGAGGGTATCTCATTTAGTAA 1691
DB 4818 CTTTACATGGTGCACAGAAATACAGAAATCTGTTGACGCTATCTTCTGCGCCACAAT 4759
QY 1692 TCAGAGCAGCAGACACAAATTCCTGGAAGCTGCAAGGCTGGAGATGT------CGA 1744
DB 4758 GCCAATCAAGCTTTAAGTACAGAGAGCGGTTTACGCACTGGCAGTTGCTTTGCAACAA 4699
QY 1745 TGTAAGAACTGTGCTACTGTTCAGAGTGTCACTGCAGAGACATTCAGAGGCGTCAGTC 1804
DB 4698 GGTACAGATGTTGGTGTGCTGCTTTGCTTTGAAATGACAGCGCGGGAAGTG-CGCTT 4640
QY 1805 TACACCACTTCAATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGTACA 1864

DB 4639 GCCAGCACTGCATATTTGCTGCTAAAGAGATGATACGAAAGAGCTAGCTATTACTTCA 4580
QY 1865 GCATGGAGCTGATGTCATGCTAAGATAAGAGAGGCTTGTACTTTGCACAATGCATG 1924
DB 4579 AAATGAGCATAACTCGGATGCTGCTCGAAAGAGGCTTTTACTCCGCTTATATCGCGC 4520
QY 1925 TTCTTATGGACATTTATGAAGTTGCAGAACTTCTTTGTTAAACATGGAGCAGTAAATGT 1984
DB 4519 TCACATATGAAATGAGAACTGAGCACTGCTACTCGAAAGAGGAGCAATGTGAATTA 4460
QY 1985 AGCTGATTTATGGAATTTACACCTTTTACATGAAGCAGCAGCAAGAAATATGAAT 2044
DB 4459 CCAAGCAGACATAACATAAGTCGTTTACACGTTTCAACAAATGGGGCTGCAACACAT 4400
QY 2045 TTGCAAACTTCTGCTCCAGCATGTTGTCAGACCTTACCAAAAAACAGGGATGAAATAC 2104
DB 4399 GGTTCGTTATTTGGTCTCATGGGCGCTAATTTGACTGTGCGACACGTGATTTACTTAAC 4340
QY 2105 TCCTTTGGATCTTTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
DB 4339 ACCATTACACTGTGCTTCTGCTTCCAGTCAATCAAGTTGTTGTTGCTTGAATAA 4280
QY 2162 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTACCCAGAGTGAAGAAGTTGCTTTC 2221
DB 4279 AGGAGCTCCAATCAGTGTCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCA 4220
QY 2222 TCCTGATATGTAATTTGCCGGATACCAAGGAGACATTCACACCTTTTACATTTAGC 2281
DB 4219 GGTGGATGATGTTACTGTGCTATCTC-----ACTCCTCTTCAATGTGGC 4175
QY 2282 AGCTGTTTATAATAATTTAGAGTTTGCAGAGTATTTTGTACAAACGAGGAGCTGATGAA 2341
DB 4174 TGCTCATTTCCGGACATGTCGCTGCTGCTAAACTTTTGTGGATCGTAACTGACCCGAA 4115
QY 2342 TCCCAAGACAAAGAGGAGCTTTTCTTTTACATATACAGCACTTACGGCATGTAGA 2401
DB 4114 TGCTCAGCTCTCAATGGCTTTCACACCTGCTATTCCTTTGCAAAAAAATTCGCAATTA 4055
QY 2402 TGTACAGCTCTACTAATAAAGTATAATGCTGCTCAATGCCAGGACAAATGGGCTTT 2461
DB 4054 AATTTGCAAGTCTTACTGAAATACCAAGCTGCAATCGAGCACTACTGATCGGTCT 3995
QY 2462 CACACCTTTTGCAGAGCAGCCCAAGAGGAGGAGCAACAGCTTTTGTGTTGCTAGC 2521
DB 3994 CTCACGCTGCTATGCTGCTGCTTTTATGGTGTCTATAAACAATGTCATCTATTACTACA 3935
QY 2522 CCATGGAGCTGACCCGACTCTTAAATAATCAGGAAGGACAAACACCTTTTAGATTTTTC 2581
DB 3934 ACAAGGTGCTAATGAGATGTTGGCTACAGTACGCGGTGAAAGCGCTCTTCATTTAGCTGC 3875
QY 2582 A 2582
DB 3874 A 3874

RESULT 11

US-09-065-474-35
; Sequence 35, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scott
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA

```

; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5235 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-065-474-35

```

Query Match 2.4%; Score 93; DB 3; Length 5235;

Best Local Similarity 45.0%; Pred. No. 6.6e-14;

Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

```

QY 1392 GGAGCAACATCAATGAAAGACTAAGAAATCTTGTACTCCTCTGACGCGGGCATCGAG 1451
D 178 GGCACCCATATACACATGCAATGCAATGGCCCTTAATGCAATGCAATGCGCCGCTCCAAA 237
QY 1452 AAAGCTCATATGATGTGTGAAGTAGTGAAGCAATGAAGCAAGAGTTAATGCTCTG 1511
D 238 GAAGGTCAATGATGAGTGTGCGGAACTCTGAAAAGAAAGGAGATGTTGATGCGCC 297
QY 1512 GATAATCTTGGTCAGACTTCTACACAGAGCTGCATATTTGTGCTCATCTACAAACCTGC 1571
D 298 ACTAGAAAGGTAACACAGCGTTACATATAGCATCTATGGCAGGACAAAGAACTAATGCTC 357
QY 1572 CGCTACTCTGAGCTATGGTGTGCTTCCATTAATATATATCCCTTCAGGGCTTTACGCT 1631
D 358 ACAGTACTTCTGAAATGCTGTAATGTTAAGCTACATCACTAAACCGGTTTACACCA 417
QY 1632 TTACAGATGGGAAATGAAATGTACAGCAACTCTCTCAAGAGGCTATCTCATTAGGTAAT 1691
D 418 CTTTACATGCTGCACAGAAATCAGAACTCTGTTGACGCTATCTCTTGGCCCAAT 477
QY 1692 TCAGAGCGACAGACAAATTTGCTGGAGCTGCAAGGCTGGAGATGT------CGAAAC 1744
D 478 GCCAATCAAGCTTTAAGTACAGAGAGCGGTTTACGCCACTGGCAGCTTGCCTTGCACAA 537
QY 1745 TGTAAGAAACCTGTACTCTTCAGAGTGTCACTGCAGAGACATTTGAAGGGCTCAGTC 1804
D 538 GGTACAGATCGTGTGGTGGCTGTTTGTGTTGAAATGACACGCGCGGAAAGTG-CGCTT 596
QY 1805 TACACCACTTCAATTTTGCAGCTGGGTATACAGAGTGTCTCGTGGTGGAAATATCTGTACA 1864
D 597 GCCAGCACTGCATATTTGCTCTAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 656
QY 1865 GCATGAGCTGATGTCATGCTTAAAGATTAAGGAGCGCTTGTACCTTTGCAATGCATG 1924
D 657 AAATGAGCATAACTCGGATGTGACTTCGAAAGCGGCTTTACTCCGCTTCATATCGCCG 716
QY 1925 TCTTATGGACATTTAAGTTTCCAGAACTCTTGTAAACATGGGAGCAGTACTTAAATGT 1984
D 717 TCNCTATGAAATGAGAACTAGCACTGCTTACTCGAAAGGGGAGCAATGTGAATTA 776
QY 1985 AGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAAAGAAATATGAAAT 2044
D 777 CCAAGCGAGACATAAATAAGTCCGTTACAGCTTGCACAAATGGGCTCGTACAAACAT 836

```

```

QY 2045 TTGCAAACTTCTGCTCCAGCATGTTGCAGACCCCTACCAAAAAACAGGATGGAATAC 2104
D 837 GGTTCGTTATGTTGGCTCATGGGCGCGTAATTGACTGTGCGCACACGTGATTTACTAAC 896
QY 2105 TCCTTTGGATCTTGTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
D 897 ACCATTACACTGTCTCTCGTTGAGTCATGATCAAGTTGTTGCTTGTGTTGCTTTGCTTGA 956
QY 2162 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAAGTTGCTTTC 2221
D 957 AGGAGCTCCAATCAGTGTCTAAGACAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 1016
QY 2222 TCCTGATAATCTAAATGCCCGCGATACCCAGGAGAGACATTCACACCTTTACATTTAGC 2281
D 1017 GGTGGATGATGTTACTGTTGACTATCTC-----ACTCCTCTTCATGTGGC 1061
QY 2282 AGCTGGTTATAATAATTTAGAAAGTTGCAGAGTATTTGTTTACACACGAGCTGATGTCAA 2341
D 1062 TGCTCAATGGGACATGTCGCTGTCGCTAACTTTTGTGCTGATCGTAATGCTGACCCGAA 1121
QY 2342 TGCCCAACACAAAGGAGGACTTATTCCTTTACATAAATGCGAGCATCTTACGGGCGATGTA 2401
D 1122 TGCTCGAGCTCTCAATGGCTTCACACGCTGCATATGCTTGCAAAAAAATCGCATTA 1181
QY 2402 TGTAGCAGCTCTACTATAAAGTATAATGCAATGTCATGTCATGTCAGGAGCAAAATGGCTTT 2461
D 1182 AATTGTCGAAGTGTCTGTAATACCGCTGCAATCGAAGCAACTACTGTAATCCGGTCT 1241
QY 2462 CACACCTTTGCAGAGCAGCCCAAGAGGAGCAACACAGCTTTGTGCTTTGTTGCTAGC 2521
D 1242 CTCACCGCTGCATGCTGCTGCTTTTATGGTGTCTATAAACATTTCTCATCTATTACTACA 1301
QY 2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
D 1302 ACAAGTGTCTAATGTCAGATGTTGGCTACAGTACGCGGTGAACGCGCTCTCTATTAGCTGC 1361
QY 2582 A 2582
D 1362 A 1362

```

RESULT 12

```

US-09-065-474-36/c
; Sequence 36, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heskia Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459

```

REFERENCE/DOCKET NUMBER: HW-5-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 5235 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-065-474-36

Query Match 2.4%; Score 93; DB 3; Length 5235;
Best Local Similarity 45.0%; Pred. No. 6.6e-14;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

QY 1392 GGAGCAACATCAATGAAAGACTAAAGAAATTCCTGACTCTCTGCACGTGGCATCTGAG 1451
DB 5058 GGCACCGATATCAACATGCAATGCGAATGCGCTTAATGCATTGCATCTGGCTCCAAA 4999
QY 1452 AAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGGTTAATGCTGTG 1511
DB 4998 GAAGTCAATCATGAAGTGGTGGCGGAATCTCTGAAAGAAAGCAGATGTTGATGCTGCC 4939
QY 1512 GATAATCTTGGTCAGACTTCTTACACAGAGCTGCATATTTGGTGCATCTACAAACCTGC 1571
DB 4938 ACTGAAAGGGTAAACACAGCTTACATATAGCATATTTGGCAGGACAAAGAACTAATCGTC 4879
QY 1572 CCTACTCTCTGAGCTATGGGTGTCATCTTAACATATATATCTCTTCAGGGGCTTTACTGCT 1631
DB 4878 ACAGTACTTGTGAAATGGTGCTAATGTTAAGCTACATCACTAAAGGTTTACACCA 4819
QY 1632 TTACAGATGGGAAATGAAATGACAGCAACTCCCTCCAAAGAGGTATCTCATATAGTAAAT 1691
DB 4818 CTTTACATGGCTGCACAAAGAAATCACGAATCTGTTGACGTATCTTCTTGGCCACAAAT 4759
QY 1692 TCAGAGGACAGACAGCAATTCGTAAGCTGCAAGGCTGGAGATGTTTTGCGCACTGGCAGTGGCTTGCACAA 4699
DB 4758 GCAATCAAGCTTTAAGTACAGAAAGACCGTTTTACGCCACTGGCAGTGGCTTGCACAA 4640
QY 1745 TGTAAAAAAGCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTC 1804
DB 4698 GGTACAGTCTGTGTGCTGCTGCTTTGTTGTTGAATGACACGCGGGAAGTG-CGCTT 4640
QY 1805 TACACCACTTTATTTGTCAGCTGGGTATTAACAGAGTGTCCGTGGTGGAAATATCTGCTACA 1864
DB 4639 GCCAGCACTGCATATGCTGCTTAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4580
QY 1865 GCATGAGCTGATGTCATGCTTAAGATAAGAGAGGCGCTTGTACCTTTGCACAATGCCATG 1924
DB 4579 AATGAGCATACTCGGATGTCGCTTGAAGAAAGCGCTTTACTCGCTTCATATCGCCGC 4520
QY 1925 TCTTTATGACATTTATGAAGTTGACAGACTTCTTGTAAACATGAGCAGTATGTTAATGT 1984
DB 4519 TCATATGAAATGAGACGTAGCACAACTGCTACTCGAAAGAGGCAATGTGAATTA 4460
QY 1985 AGCTGATTTATGAAATTTACACCTTTTACATGAAGCAGCAGCAAAAGAAATATGAAT 2044
DB 4459 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGTCGTACAAACAT 4400
QY 2045 TTCCAAACTCTCTCCAGCATGGTGCAGACCTACCAAAAAAACAGAGGATGGAATAC 2104
DB 4399 GGTTCGTTATTTGGTCTAGTGGGCGCGTAATTTGACTGTGCGACACGTGATTTACTAC 4340
QY 2105 TCCTTTGGATCTTTTAAAGATGGAGATACAGAT---ATTCAAGATCTCTTTAGGGGAGA 2161
DB 4339 ACCATTACACTGTGCTCTCTGTTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGA 4280
QY 2162 TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGAGTGAAAGTTGCTTC 2221
DB 4279 AGGAGCTCCCAATCAGTCTAAGACAAAAAATGGTTGGCTCCCTTACATATGCGACACA 4220

QY 2222 TCCTGATATGTAAATGCGCGATACCCAGGACAGACATTTCAACACCTTTTACATTTAGC 2281
DB 4219 GGTGATGATGTTACTGTTGACTATCTC-----ACTCCTCTTTCATGTGGC 4175
QY 2282 AGCTGTTTATAATAAATTTAGAGTTGACAGATATTTGTTTACACAGGAGCTGATGTGAA 2341
DB 4174 TGTCTATTGGGACATGTCGCTGCTAAACTTTTGTGCTGATCGTAAATGCTGACCCGAA 4115
QY 2342 TGCCCAAGACAAAGGAGACTTATCTTTACATATGAGCATCTTACGGGCGCATGTAGA 2401
DB 4114 TGCTGAGCTCTCAATGGCTTCACACCGTGCATATCGCTTGCAAAAAATTCGCATTA 4055
QY 2402 TGTCAGAGCTCTACTAATAAAGTATATGATGCTGTCAATGCCACGACAAATGGCTTT 2461
DB 4054 AATGTCGAAGCTCTACTGAAATACCACTGCAATCGAAGCAACTACTGAAATCCGCTCT 3995
QY 2462 CACACCTTTGACGAGACGCCCAAAAGGACGACACACAGCTTTGCTTTGCTGCTAGC 2521
DB 3994 CTCACGCTGCTATGCTGCTCTTTTATGGTGTCTATTAACATTTGTCATCTATTACTACA 3935
QY 2522 CCATGAGCTGACCGGACTCTTAAAAATCAGGAAGGACGACAAACACCTTTAGATTTAGTTTC 2581
DB 3934 ACAAGGTGCTAATGCAGATGCTGCTACAGTACGCGGTGAACGCCCTCTTCATTTAGCTGC 3875
QY 2582 A 2582
DB 3874 A 3874

RESULT 13
US-09-557-034-35
; Sequence 35, Application US/09557034
; Patent No. 6365569
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; PROTEINS, NUCLEIC ACID MOLECULES, AND
; USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,034
; FILING DATE: 21-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065,474
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5235 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

Db 5058 GCACCGATATCAACACATGCAATGGAATGCGCTTAATGTCATGTCGCTCCAAA 4999
QY 1452 AAGCTCATTAATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTG 1511
Db 4998 GAAGGTGTCATGATGAAGTGGTCCGCAAACTTCGAAAAGAAAGACAGATGTTGATGCTGC 4939
QY 1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATGTTGGTCATCTACAAACCTGC 1571
Db 4938 ACTAGAAAGGTTAACACAGCGTTACATATAGCATCATGCGCAGGACAACTAATCGTC 4879
QY 1572 CGCCTACTCTGAGCTATGGGTGTGATCCTTAACATTAATATCCCTTCAGGGCTTTACTGCT 1631
Db 4878 ACAGTACTTGTGAAAATGCTGCTAATGTTAAGTACAACTCACTAAACGGTTTACACCA 4819
QY 1632 TTACAGATGGGAATGAAAATGTACAGCAACTCTCCCAAGAGGTTATCTCATTAGTAA 1691
Db 4818 CTTTACATGGCTGCACAGAAAATCAGCAATCTGTTGACGCTATCTTCTGCCCAAT 4759
QY 1692 TCAGAGGCACAGACAAATGCTGGAAGCTGCAAGGCTGGAGATGT-----CGAAAC 1744
Db 4758 GCAATCAAGCTTTAAGTACAGAACGCTTTACGCCACTGCGATTTGCCCTTGCAACAA 4699
QY 1745 TGTAAAAAAGCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTC 1804
Db 4698 GGTCAAGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4640
QY 1805 TACACCACTTCAATTTGCAAGCTGGGTATTAACAGAGTGTCCGTGGTGGATATCTGCTACA 1864
Db 4639 GCGACACTGCAATATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4580
QY 1865 GCATGAGCTGATGTCATGCTAAAGATAAAGGAGCGCTTGTACCTTTGCACAAATGCATG 1924
Db 4579 AATGAGCATACTCGGATGTGACTTCGAAAAGCGCTTTACTCCCTTCATATCGCCGC 4520
QY 1925 TTCTTATGACATTAATGAAGTTCGAGAATCTTTGTTAAACATGAGCAGTAGTAAATG 1984
Db 4519 TCCTATGAAATGAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4460
QY 1985 AGCTGATTTATGGAATTTACACCTTTACATGAGCAGCAGCAAGAAAGAAATATGAAAT 2044
Db 4459 CCAAGCGACACATACATGAAGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4400
QY 2045 TTGCAAACTTCTGCTCGACATGGTGCAGACCTACCAAAAAAAGAGGATGGAATAC 2104
Db 4399 GGTTCGTTTATGTTGGCTCATGGGCGCTAATGACTGTGCACACGATGATTTACTAAC 4340
QY 2105 TCCTTTGGATCTGTTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
Db 4339 ACCATTACACTGTGCTTCTGCTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTCAAAA 4280
QY 2162 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAGTTGCTTTC 2221
Db 4279 AGGAGCTCCAATCAGTGTGAAGCAAAAAATGTTGGCTTCCCTTACATATGCGCAGACA 4220
QY 2222 TCCTGATAATGTAATGTCGGGATACCCCAAGCAGACATTCACACCTTTTACATTTAGC 2281
Db 4219 GGTGGATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4175
QY 2282 AGCTGTTTATAATTAATTAAGATGTCAGAGTATTTGTTTACACACGGAGCTGATGAA 2341
Db 4174 TGTCTATTGGGACATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4115
QY 2342 TGCCCAAGACAAGGAGGACTTATTCCTTTACATATGAGCAGCATCTTACGGGATGTAGA 2401
Db 4114 TGTCCAGCTCTCAATGTTGGCTTACACCGCTGTCATATCGCTTGCACAAAAATGCGATTAA 4055
QY 2402 TGTAGCAGCTCTACTAATAAGTATATGATGTGTGATGTCACGACGACAAATGGGCTTT 2461
Db 4054 AATTGTGCACTGCTACTGAATACACCGTGAATTCGAGCAACACTGATGAATCCGGTCT 3995
QY 2462 CACACCTTTGCAGGAAGCGCCCAAAAGGACGACACAGCTTTGTGCTTTGTGCTAGC 2521

Db 3994 CTCACGGCTGCAATGCTGCTCTTTTATGGTGTCTATAAACAATTCATCTATTTACTACA 3935
QY 2522 CCATGAGCTGACCGACTCTTAAAAATCAGGAAGCAACACACCTTTAGATTTAGTTTC 2581
Db 3934 ACAAGTGTCTAATGCAAGATGTGGCTACAGTACCGGTGAAACGCCCTTTCATTTAGCTGC 3875
QY 2582 A 2582
Db 3874 A 3874
RESULT 15
US-09-031-485-32
; Sequence 32, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIOPHILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Hesk Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5503 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51..5285
; US-09-031-485-32

Query Match 2.4%; Score 93; DB 1; Length 5503;

Best Local Similarity 45.0%; Pred. No. 6.8e-14;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

QY 1392 GGAGCAACATCAATGAAAGACTTCTGACTCTCTGACGGTGGCATCTGAG 1451
Db 228 GGCACCGATATCAACACATGCAATGCGAATGGCTTAATGATTCGCTGCCCAAA 287
QY 1452 AAGCTCATTAATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTG 1511
Db 288 GAAGTCAATCATGAAGTGGTCCCGCAACTTCTGAAAGAAAAGCAGATGTTGATGCTGCC 347

Qy	1512	GATTAATCTTGGTCAGACTTCTCTCACAGAGCTGCAATATGTGGTCATCTACAAACCTGC	1571
Db	348	ACTAGAAGGTTAACACAGCGTTTACATATAGCATCATTTGGCAGGACAGAACTAATCGTC	407
Qy	1572	CGCTACTCTCGAGCTATGGGTGTGATCCTAACATTTATATCCCTTCAGGCGTTTACTGCT	1631
Db	408	ACAGTACTTGTTCAAAATGGTGTATGTTACGTACAAATCACAACGGTTTTTACACCA	467
Qy	1632	TTACAGATGGGAATGAAATGTACAGCAACTCTCCCAAGAGGGTATCTCATTAGGTAAAT	1691
Db	468	CTTTTACATGGCTGCACAAGAAAATCAGGAATCTGTGTGTACGCTATCTCTTGCGCCACAAT	527
Qy	1692	TCAGAGCGACAGACAGCAATTTGCTGGNAGCTGCAAGGCTGGAGATGT-----CGAAAC	1744
Db	528	GCCNATCAAGCTTTAGTCAGAAGACGGTTTTTACGCCACTTGGCAGTTGCTCTGCAACAA	587
Qy	1745	TGTAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGCAAGACACATTTGAAGGGGTGAGTC	1804
Db	588	GGTCACGATCGTGTGTCGCTGTTTTGCTTGAAATGACACGCGCGGAAGTG-CGCTT	646
Qy	1805	TACACACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAAATATCTGTACA	1864
Db	647	GCCAGCACTGCATTTTGTGCTGTAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA	706
Qy	1865	GCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGCGCTTGTACCTTTGTCACAATGCATG	1924
Db	707	AAATGACCATTAATTCGGATGTGACTTCGNAAGCGCTTTTACTCCGCTTCATATCGCCGC	766
Qy	1925	TTCTTTAGACATTTAAGTTTGCAGAACTTCTTTTAAACATGGACAGTAGTTAATGT	1984
Db	767	TCACTATGGAATGAGAAGCTAGCAACTGCTACTCGAAAGGAGCAACTGTGAATTA	826
Qy	1985	AGCTGATTTATGAAATTTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT	2044
Db	827	CCAAGCGAGACATAACATAAGTCCGTTACCGTTGCACAAAAATGGGGTCGTACAACAT	886
Qy	2045	TTGCCAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAACAGGATGGAATAC	2104
Db	887	GGTTTCGTTATTCTTGGCTCATGGGCGCTAATTGACTGTGCGACACGTGATTTACTTAAC	946
Qy	2105	TCCTTTGGATCTTGTTTAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGAGA	2161
Db	947	ACCATTACATGTGCTTCTGTTTCAGGTCATGATCAAGTTGTTGATTTGTGCTGAAAA	1006
Qy	2162	TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTAGCCAGAGTGGAAGTTGTCTTC	2221
Db	1007	AGGAGCTCCAATCAGTGCTAAGACAAAAATGGTTTGCTCCCTTACATATGCGACACA	1066
Qy	2222	TCCTGATATGTAAATTTGCCGGCATACCAAGCGAGACATTCACACCTTTACATTTAGC	2281
Db	1067	GGTGGATGATGTTACTGTTGCACTATCTC-----ACTCCTCTTCATGTGC	1111
Qy	2282	AGCTGGTTATATAATTTAGAAGTTGCAGAGTATTTTGTTCACACAGGCGATGTGAA	2341
Db	1112	TGCTCATTCGGGACATGCTCGTGTCGTAAACTTTTGTGATGCTGAATGCTGACCCGAA	1171
Qy	2342	TGCCAAGACAAAGGAGACTTTTTCCTTTTACATAATGCAGCACTTTTACGGGCGATAGA	2401
Db	1172	TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATGTCGTTGCAAAAAAATCGCATTA	1231
Qy	2402	TGTAGAGCTCTACTAATAAGATATAATGCATGTCTCAATGCCAGGACAAATGGGCTTT	2461
Db	1232	AATTGTGAACTGTACTGAAATACCGCTGCAATCGAGCAACTACTGAATCCGGGTCT	1291
Qy	2462	CACACCTTTGCAGAGCAGCCGCAAAAGGAGACACAGCTTTTGTCTTTTGTGTAGC	2521
Db	1292	CTCACGCTGCAATGTCGCTGTTTTATGGTGCTATAAAACATTGTCATCTATTACTACA	1351
Qy	2522	CCATGAGCTGACCCGACTCTTTAAAAATACAGAAGGACAAACCTTTGATTTAGTTTC	2581
Db	1352	ACAAGTGCTTAATGACAGATGTGGCTACAGTACGGGGTGAACGCGCTCTCATTTAGCTGC	1411
Qy	2582	A	2582

Db 181 CGAAATACAGATGGAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 240
QY 741 ACTGCTGAATATAAGAAAGATGAACCTCTAGAAAGTGCAGGAGTGGCAATGAAGAAAA 800
Db 241 ACTGGTGAATATAAGAAAGATGAACCTCTAGAAAGTGCAGGAGTGGCAATGAAGAAAA 300
QY 801 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 860
Db 301 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 360
QY 861 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGATTTGACAGCTGTACTGCA 920
Db 361 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGATTTGACAGCTGTACTGCA 420
QY 921 CATGAGCTGATGTCATCTCAAGATAAAGTGTATCTGGTACCATTACACAAATGCCTGT 980
Db 421 CATGAGCTGATGTCATCTCAAGATAAAGTGTATCTGGTACCATTACACAAATGCCTGT 480
QY 981 TCTTATGGTCAATTAAGATAAAGTGTATCTGGTACCATTACACAAATGCCTGT 1040
Db 481 TCTTATGGTCAATTAAGATAAAGTGTATCTGGTACCATTACACAAATGCCTGT 540
QY 1041 ATGGACTTGTGGCAATTCACCTCTTATGAGGCGAGCTTCTAAGAACAGGCTTGAAGTA 1100
Db 541 ATGGACTTGTGGCAATTCACCTCTTATGAGGCGAGCTTCTAAGAACAGGCTTGAAGTA 600
QY 1101 TGTCTCTCTCTTAAGTATGTTGTCAGACCCCAACACTGCTCAATTGTGCACAAATAAAGT 1160
Db 601 TGTCTCTCTCTTAAGTATGTTGTCAGACCCCAACACTGCTCAATTGTGAAGATAAAGT 660
QY 1161 GGTATAGACTTGGCTGCCACACAGTGTAAAGAAAGATTAGCATATAAATTTAAAGGC 1220
Db 661 GGTATAGACTTGGCTGCCACACAGTGTAAAGAAAGATTAGCATATAAATTTAAAGGC 720
QY 1221 CACTCGTGTCTCAAGCTGCAGGAGAGTGTACTCGAATCAAAAACATCTCTCT 1280
Db 721 CACTCGTGTCTCAAGCTGCAGGAGAGTGTACTCGAATCAAAAACATCTCTCT 780
QY 1281 CTGGAATAGGTGAATTTCAAGCATCTCAACACATGAACAGCATTTGCTGTCTGTCT 1340
Db 781 CTGGAATAGGTGAATTTCAAGCATCTCAACACATGAACAGCATTTGCTGTCTGTCT 840
QY 1341 GCATCTCCATATCCAAAGAAAGCAATATGTGAACCTTGTCTAAGAAAGGAGCAAAAC 1400
Db 841 GCATCTCCATATCCAAAGAAAGCAATATGTGAACCTTGTCTAAGAAAGGAGCAAAAC 900
QY 1401 ATCAATGAAAGACTAAAGAAATCTTGACTCTCTGCAGTGGCATCTGAGAAAGCTCAT 1460
Db 901 ATCAATGAAAGACTAAAGAAATCTTGACTCTCTGCAGTGGCATCTGAGAAAGCTCAT 960
QY 1461 AATGATGTTGTTGAAGTGTGTTGAACATGAAGCAAAAGGTTAATGCTCTGGATTAATCTT 1520
Db 961 AATGATGTTGTTGAAGTGTGTTGAACATGAAGCAAAAGGTTAATGCTCTGGATTAATCTT 1020
QY 1521 GGTACAGCTTCTACAGAGCTGCATATTTGTGTCATCTCAAAACCTGCGCGCTACTC 1580
Db 1021 GGTACAGCTTCTACAGAGCTGCATATTTGTGTCATCTCAAAACCTGCGCGCTACTC 1080
QY 1581 CTGAGCTATGGGTGTGATCCTAACATTAATCCCTTCAGGGCTTTACTGCTTTACAGATG 1640
Db 1081 CTGAGCTATGGGTGTGATCCTAACATTAATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140
QY 1641 GGAATGAAATGTACAGAACTCTCCAAAGAGGTATCTAATAGTAATTCAGAGGCA 1700
Db 1141 GGAATGAAATGTACAGAACTCTCCAAAGAGGTATCTAATAGTAATTCAGAGGCA 1200
QY 1701 GACAGCAATTCCTGGAAGCTGCAAGGCTGAGATGTGCAAACTGTAAAAAACTGTGT 1760
Db 1201 GACAGCAATTCCTGGAAGCTGCAAGGCTGAGATGTGCAAACTGTAAAAAACTGTGT 1260
QY 1761 ACTGTTACAGAGTGTCAACTGACAGACATTTGAAGGCGCTCAGTCTACACCACTTCATTTT 1820
Db 1261 ACTGTTACAGAGTGTCAACTGACAGACATTTGAAGGCGCTCAGTCTACACCACTTCATTTT 1320

QY 1821 GCAGCTGGTATAACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTG 1880
Db 1321 GCAGCTGGTATAACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTG 1380
QY 1881 CATGCTAAAGATAAAGAGGCGCTTGTACCTTTGCGCAATGCAATGCTTTATGGAATTTAT 1940
Db 1381 CATGCTAAAGATAAAGAGGCGCTTGTACCTTTGCGCAATGCAATGCTTTATGGAATTTAT 1440
QY 1941 GAAAGTGGAGAACTCTCTTAAACATCGACAGTAGTTAATGTAGCTGATTTATGGAAA 2000
Db 1441 GAAAGTGGAGAACTCTCTTAAACATCGACAGTAGTTAATGTAGCTGATTTATGGAAA 1500
QY 2001 TTTTACACCTTTACATGAGCAGCAGCAAAAGAAATATGAAATTTGCAAACTCTTGCTC 2060
Db 1501 TTTTACACCTTTACATGAGCAGCAGCAAAAGAAATATGAAATTTGCAAACTCTTGCTC 1560
QY 2061 CAGCATGGTGGCAGACCCCTACCAAAAACACAGGATGGAATACTCCTTTGGATCTTGT 2120
Db 1561 CAGCATGGTGGCAGACCCCTACCAAAAACACAGGATGGAATACTCCTTTGGATCTTGT 1620
QY 2121 AAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCT 2180
Db 1621 AAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCT 1680
QY 2181 GCCAAGAAAGGTTGTTAGCCAGAGTGAAGAAAGTTGCTCTCTCTCATATGTAATTTGCT 2240
Db 1681 GCCAAGAAAGGTTGTTAGCCAGAGTGAAGAAAGTTGCTCTCTCTCATATGTAATTTGCT 1740
QY 2241 CCGCATACCAAGGAGCAGCATTTCAACACCTTTACATTTAGCAGCTGGTTATATATTTA 2300
Db 1741 CCGCATACCAAGGAGCAGCATTTCAACACCTTTACATTTAGCAGCTGGTTATATATTTA 1800
QY 2301 GAAAGTGGCAGATATTTGTTTACAACAGGAGCTGATGTAATGCCCAAGCAAAAGGAGCA 2360
Db 1801 GAAAGTGGCAGATATTTGTTTACAACAGGAGCTGATGTAATGCCCAAGCAAAAGGAGCA 1860
QY 2361 CTTATTTCTTTACATAATGACAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2420
Db 1861 CTTATTTCTTTACATAATGACAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920
QY 2421 AAGTATATGCAATGTCAANTGCCAGGCAAAATGGCTTTCCACACCTTTGCAAGCAAGCA 2480
Db 1921 AAGTATATGCAATGTCAANTGCCAGGCAAAATGGCTTTCCACACCTTTGCAAGCAAGCA 1980
QY 2481 GCCCAAAAGGAGCAGCAACAGCTTTGCTTTGTTGCTAGCCCATGAGCTGACCCGACT 2540
Db 1981 GCCCAAAAGGAGCAGCAACAGCTTTGCTTTGTTGCTAGCCCATGAGCTGACCCGACT 2040
QY 2541 CTTTAAATTCAGGAAGGCAAAACACCTTTTAGATTTAGTTTTCAGCGGATGATGTCAGCGCT 2600
Db 2041 CTTTAAATTCAGGAAGGCAAAACACCTTTTAGATTTAGTTTTCAGCGGATGATGTCAGCGCT 2100
QY 2601 CTTCTTGACAGCAGCCATGCCCATCTGCTCTGCCCTCTTGTACAAAGCTCAAGTGTCTC 2660
Db 2101 CTTCTTGACAGCAGCCATGCCCATCTGCTCTGCCCTCTTGTACAAAGCTCAAGTGTCTC 2160
QY 2661 AATGTTGTGAGAAGCCCGAGGAGCCACTGTCAGATGCTCTCTCTTCAGGTCATCTAGCCCA 2720
Db 2161 AATGTTGTGAGAAGCCCGAGGAGCCACTGTCAGATGCTCTCTCTTCAGGTCATCTAGCCCA 2220
QY 2721 TCAGCGCTTTCTGACGCCAGCAGTCTTGACAACCTTTATCTGGGAGTTTTTCAGAACTGTCT 2780
Db 2221 TCAGCGCTTTCTGACGCCAGCAGTCTTGACAACCTTTATCTGGGAGTTTTTCAGAACTGTCT 2280
QY 2781 TCAGTATGTTAGTTCAAGTGAACAGAGGCTGCTCCAGTTTGGAGAAAAGAGGTTCCA 2840
Db 2281 TCAGTATGTTAGTTCAAGTGAACAGAGGCTGCTCCAGTTTGGAGAAAAGAGGTTCCA 2340
QY 2841 GGAGTAGATTTTAGCATAACTCAATTCGTAAGAAATCTTTGGACTTTGAGCACCTTAATGGAT 2900
Db 2341 GGAGTAGATTTTAGCATAACTCAATTCGTAAGAAATCTTTGGACTTTGAGCACCTTAATGGAT 2400

Db 927 CCAAAACATTGGGAACACTGATGGGAAATCAGCCCTGGCACTGGCAGATCCTTTACAGAAAA 986
Qy 732 GCAGTCCCTTACTGCTCAATATTAAGAAAGATGAACCTTTAGAAGTCCAGAGAGTGGCAAT 791
Db 987 GCTGCTCCTTACAGGTGAATACAAAGAAAGACAGAACCTTAGAAGCTGCTAGGAGTGGTAAT 1046
Qy 792 GAAGAAAAATGATGGCTCTACTACACCAATTAATGTCAACTGCCACCAAGTATGGC 851
Db 1047 GAAGAAAAACTAATGGCTTTACTGACTCTCTTAATGTGAATGGCATGCCATGCCAGTATGGG 1106
Qy 852 AGAAGTCAACGCCAATACATTTGGCAGCAGAGATATAACAGAGTAAGAATGTACAGCTG 911
Db 1107 CGAAAGTCACCTCTTACATCTAGCAGGGGCTACACAGAGTTCCGAATAGTTCAGCTT 1166
Qy 912 TTACTGCAACATGGAGCTGATGTCATGCTAAAGATAAAGGTGATCTGCTACCATTTACAC 971
Db 1167 CTTCTTCAGCATGGTGTGATGTTATGCAAAAGACAAAGGTGGAGCTTGCTGCTCTTCAT 1226
Qy 972 AATGCTGTTCTTATGCTCATTTAAGTAAGTGAACCTTTTGGTCAAGCATGGTGCCTG 1031
Db 1227 AATGCTGTTCTATGACATATTAAGTACAGAACTGCTACTAAAGCATGGAGCTTGT 1286
Qy 1032 GTAATGCAATGGACTTGTGGCAATTCATCTCTTCATGAGCAGCTTCTTAAGAACAGG 1091
Db 1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACAGGGCTGCTTCCAAAGAACGT 1346
Qy 1092 GTTGAAGTATGTTCTCTTCTTAAAGTTATGTTGTCAGACCCCAACACTGCTCAATTGTCAC 1151
Db 1347 GTAGAAGTCTGCTTGTGTTACTTAGCCATGGCGTGATCTTACGTTAGTCAACTGCCAT 1406
Qy 1152 AATAAAGTGCATAGACTTGGTCCACACCACAGCTTAAAGAAAGATTAAGCATATGAA 1211
Db 1407 GGCAAAAGTGTGTGATATGGCTCCAACTCCGGAGCTTAGGAGAGATTGACTTTATGAA 1466
Qy 1212 TTAAAGGCCACTGTTGCTGCAAGCTGCAGNAGAGCTGATGTTACTCGAATCAAAAA 1271
Db 1467 TTTAAAGGTCATCTTTACTACAGCAGCCAGAGAGAGACTTAGCTAAAGTTAAAAAA 1526
Qy 1272 CATCTCTCTCGGAATGTGTAATTTCAAGCATCTTCAACACATGAAACAGCATTTGCAT 1331
Db 1527 ACATCGCTCTGGAATCATTAATTTCAACACAGCCAGCTCTCATGAAACAGCACTGCAC 1586
Qy 1332 TGTGCTGCTGCACTCCATATCCAAAAGAAAGCAAAATATGTAAGTGTGCTAAGAAAA 1391
Db 1587 TGTGCTGTGGCCTCTGCTGATCCCAACCTAAACAAAGTACAGAAATGTTACTTAGAAAA 1646
Qy 1392 GGAGCAACATCAATGAAAGACTAAGAAATTTCTTCACTCTCTGACAGTGGCATCTGAG 1451
Db 1647 GGAGCAAAATGTTAAGAAAAAATAAAGATTTTCACTGCTCCCTGCTGATGTCAGCCGAA 1706
Qy 1452 AAAGCTCAATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTG 1511
Db 1707 AGAGCCCAATATGATGTCATGGAAGTTCTGCTAATGATGCGCCCAAGATGAATGCACTG 1766
Qy 1512 GATAATCTTGGTCAGACTTCTTACACAGAGTGCATATTTGGTGCATCTAACAACCTGC 1571
Db 1767 GACACCTTGGTCAGACTCTTGTGATAGAGCGCCTAGCAGGCCACCTGCGAGACCTGC 1826
Qy 1572 CGCTACTCTGAGCTATGGTGTGATCTTAACATTAATCTTCCCTTACAGGCTTTACTGCT 1631
Db 1827 CGCTCTCTGCTGAGTTACGGCTGTGACCCCTCCATCATCTCTTACAGGCTTACAGCA 1886
Qy 1632 TTACAGATGGGAATGAAATGTACAGCAACTCTCTCAAGAGGGTATCTCATTAGGTAAT 1691
Db 1887 GCACAGATGGCAATGAACAGTGCAGCAGATTTCTGAGTGAGAGTACACCTATACGTACT 1946
Qy 1692 TCAGAGGCACAGACAATTTGCTGGAGCTGCAAGGCTGGAGATGTCGAAACTGTAAAA 1751
Db 1947 TCTGATGTTGATTATCGACTCTTAGAGGATCTTAAAGCTGGAGCTTGGAAACTGTGAAG 2006
Qy 1752 AAACCTGTGACTTTCAGAGTGTCACTGCAGACATTTGAAGGGGCTCAGTCTACACCA 1811
Db 2007 CAACTTTGAGCTCTCAAAATGTGAATTTAGAGACTTTAGAGGGCCGGCATTTCCACGCC 2066

Qy 1812 CTTCAITTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGA 1871
Db 2067 TTACACTTCCGACAGGCTTACACCGCTGTCTGTGTAGAGTACCTGCTACACACGGT 2126
Qy 1872 GCTGATGTCATGCTAAAGATAAAGAGAGCCCTTGTACCTTTGCACAATGCATGTTCTTAT 1931
Db 2127 GCCGATGCTCATGCTCAAGAACAGAGGCTGCTGGTCCCTTCATATATGCTGTTTCATAT 2186
Qy 1932 GCACATTAAGAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTAATGTAGCTGAT 1991
Db 2187 GGACACTATGAGGTGGCTGAGCTTTTAGTAAGCATGGGCTTCTCTAATGTGGCGAC 2246
Qy 1992 TTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAAGGAAAAATGAAATTTTCCAA 2051
Db 2247 TTATGGAATTTTACCCCTCTCCATGAAGCAGCAGCTTAAGAAAGATGTAATCTCAAG 2306
Qy 2052 CTTCTGCTCCAGCATGTCAGACCCCTACCAAAAAAAGAGGATGAAATACTCCTTTG 2111
Db 2307 CTCCTTTTAAACATGGAGCAGATCCAACCTAAAAAGACAGAGATGAAATACACCTTTG 2366
Qy 2112 GATCTGTTAAAGATGGAGATACAGATATCAAGATCTGCTTAGGGGAGATGCAGCTTG 2171
Db 2367 GATTTGGTAAAGAGGAGCAGACAGATATTCAGGACTTACTGAAGGGGATGCTGCTTG 2426
Qy 2172 CTAGATGCTGCCAAAGAGGCTTTTAGCCAGAGTGAAGAAGTTGCTTCTCTGTATAAT 2231
Db 2427 TTGGATGCTGCCAAGAGGCTGCCCTGGCAGAGTGCAGAAGCTCTGTACCCAGAGAA 2486
Qy 2232 GTAATTTGGCGGATACCCAAAGCAGACATTCACACCTTTTACATTTAGCAGCTGTTAT 2291
Db 2487 ATCACTGCGAGAGACACCCAGGGCAGAAATTCACCCCTCTGCACCTGGCAGCAGCTAT 2546
Qy 2292 AATAATTTAGAAGTTCAGAGTATTTGTACACACGAGAGTGTGATGATGCCAAGAC 2351
Db 2547 AATACTGGAAGTAGCTGAATATCTTAGAGCATGGAGCTGATGTTAATGCCCAGAG 2606
Qy 2352 AAAGGAGGACTTATTCCTTTTACATAATGAGCATCTTAGGGCATAGATGTAGCAGCT 2411
Db 2607 AAGGGTGGTTAATCTCTTTCATATGCGGATCTTATGGCATGTTGACATAGCGCT 2666
Qy 2412 CTACTAATAAGTATTAATGATGTGTCATGCCACGGCAAAATGGCTTTCACACCTTG 2471
Db 2667 TTATGATATAATACACACGCTGTGTAATGCAACAGATAAGTGGGCTTTACTCCCTC 2726
Qy 2472 CACGAAGCAGCCAAAAGGAGCAACACAGCTTTGTGCTTGTAGTCCCATGGAGCT 2531
Db 2727 CATGAAGCAGCCAGAAAGAGGAGCGAGCTGTGCGCCCTCTCTCTAGGCAATGGCA 2786
Qy 2532 GACCCGACTCTTAAAAATCAGGAAGCAACACCTTTAGATTTAGTTTTCAGCGGATGAT 2591
Db 2787 GACCCCACTGAAGAACCCAGGAAGCCAGACGCTCTGATCTGGCAACAGCTGACGAT 2846
Qy 2592 GTCAGGCTCTTCTGACACAGCCATGCCCCATCTGCTGCTGCCCTCTTGTTCACAGCT 2651
Db 2847 ATCAGAGCTTTGCTGATAGATGCCATGCCCGCAGAGGCTTACCTACCTGTTTAAACCT 2906
Qy 2652 CAAGTCTCAATGCTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTTCCAGTCCA 2711
Db 2907 CAGG-----CTACTGTAGTGAAGTGGCTCTCTGATCTCACAGCA 2945
Qy 2712 TCTAGCCCATCAAGCCTTTCTGACGAGCAGAGCTTTGACAACCTTATCTGGGAGTTTCA 2771
Db 2946 TCCACCCCTCTCTGCTCTCGGCTGCCAGCAGCATAGACAACCTCTACCTGSCCTTTAGCA 3005
Qy 2772 GAAGTCTCTCAGTAGTTAGTTCAGTGAAACAGAGGGTCTTCCAGTTTGGAGAAAAAG 2831
Db 3006 GAGTTGGCGCTAGGAGAGCCTCCAATGAGGGGATGGCGCGCGGAAACAGAAAGAAAG 3065
Qy 2832 -----GAGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGAACTCTTGGACTT 2885
Db 3066 GAAGGAGAAGTGTCTGTCTTGACATGAATATCAGCCATTTCTAAAAAGCCTTGGCCTT 3125

QY 792 GAAGAAAAATGATGGCTTCTACTCACACCAATAAATGTCAAACGCCAGCGAATGATGGC 851
D 1047 GAAGAAAAAATAATGGCTTTACTGACTCTCTAAATGTGAATGCCATGCCAAGTGAATGGG 1106
QY 852 AGAAGTCAACTCCATACATTTGGCAGCAGGATATACAGAGTAAAGATGTACAGCTG 911
D 1107 CGAAGTCCGACTCTTTACATCTAGCAGCGGCTTACAACAGAGTTCGAATAGTTCAGCTT 1166
QY 912 TTACTGCAACATGAGCTGATGCTCAATGCTAAAGATAAAGTGAATGCTGGTACCATTACAC 971
D 1167 CTTCCTTCAGCATGGTGTGATGCTTATGCAAAAGACAAAGGTGGACTGTGCCCTCTTCAT 1226
QY 972 AATGCCCTGTTCTTATGCTCAATTAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
D 1227 AATGCATGTTTCATATGACATTTATGAAGTACAGAACTGCTACTAAGCATGGAGCTTGT 1286
QY 1032 GTAATCCAAATGGACTGTGGCAATTCACCTCTTTCATGAGGAGCTTCTTAAGAACAGG 1091
D 1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCTG 1346
QY 1092 GTTGAAGTATGTTCTCTCTTAAAGTTATGGTGCAGACCCCAACACTGCTCAATGTTCAC 1151
D 1347 GTAGAAGTCTGCTTTTACTTAGCCATGGCCTGATCCTACGTTAGTCAACTGGCAT 1406
QY 1152 AATAAAGTCTATAGACTTGGCTCCACACCACAGTTAAAGAAAGATAGCATATGAA 1211
D 1407 GGCAAAAGTCTGTGGATATGGCTCCAACCTCGGAGCTTAGGGAGAGATTGACTTATGAA 1466
QY 1212 TTTTAAAGGCCACTGTTGCTGCAAGCTGCACGAGAGCTGATGTTACTCGAATCAAAAA 1271
D 1467 TTTTAAAGTCAATCTTTACTACAAAGCAGCAGAGAGACTTAGCTAAAGTTAAAAA 1526
QY 1272 CATCTCTCTCGAAATGGTGAATTTCAAGCATCCTCAACACATGAACAGCATGGAT 1331
D 1527 ACATCGCTCTGGAATCAATTAATCAAAACACCCAGCTCTCATGAACAGCACTGCAC 1586
QY 1332 TGTGCTGCTGCATCTCCATATCCAAAAGAAAGCAATATGTGAATGTTGCTTAAGAAA 1391
D 1587 TGTGCTGTGGCTCTCTGCTATCCCAACGCTTAAACAGTACAGAAATGTTACTTAGAAA 1646
QY 1392 GGAGCAACATCAATGAAAGACTAAGAAATCTTGACTCCTCTGCACGTTGGCATCTGAG 1451
D 1647 GGAGCAAAATGTTAATGAAAAAATAAAGATTTCATGACTCCCTGCTGATGTTGCAGCCGA 1706
QY 1452 AAAGCTCAATAAGTATGTTGAGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTG 1511
D 1707 AGAGCCCATTAATGATGCTATGGAAGTTCTGCAATAGCATGGGCCAAGATGAATGCACTG 1766
QY 1512 GATAATCTTTGGTCAGACTTCTTACACAGAGCTGCATATTTGGTCATCTACAAACCTGC 1571
D 1767 GACACCTTTGGTCAGACTGCTTTGCATAGAGCCGCTTAGCAGGCCACCTGCAGACTGC 1826
QY 1572 CGCCTACTCCTGAGCTATGGGTGTGATCTCAATATATATCCCTTCAGGGCTTTACTGCT 1631
D 1827 CGCCTCCTGCTGAGTTACGGCTCTGACCCCTCCATCATCTCCTTACAAGGCTTCACAGA 1886
QY 1632 TTACAGATGGGAATCAAAATGTACAGAACT----- 1663
D 1887 GCACAGATGGGCAATGAAGCAGTGCAGCAGATTCTGAGTGTGAGTTACGGCTCTGACCCC 1946
QY 1664 ----- 1663
D 1947 TCCATCATCTCCTTACAGGCTTTCACAGCAGCAGATGGGCAATGAAGCAGTGCAGCAG 2006
QY 1664 ----- 1663
D 2007 ATTCTGAGTGTTCATTCGTAGATAGTATCTACTTCTACCCCTTAATGGTGTATCTGA 2066
QY 1664 ----- 1663
D 2067 GACGGGAAGATTTAGAAGAAATCTATCCAGCATGCTTTCATGTCACATGTCAACATGAAGAGTAC 2126

QY 1664 ----- 1663
D 2127 ACCTATACGTACTTCTGTATGTTGATTATCGACTCTTTAGAGGCATCTAAAGCTGGAGACTT 2186
QY 1664 ----- 1663
D 2187 GGAACCTGTGAAGCAACTTTGCAGCTCTCAAAATGTGAATTTGTAGAGACTTTAGAGGGCGC 2246
QY 1664 ----- CCTCCAAGAGGCTATCTCATTTAGGTAATTC 1694
D 2247 GCATTCACGCGCTTACACTTCGCAGCAGGCTACAACAGAGTACACCTATAGTACTTCT 2306
QY 1695 GAGCAGACACACAATTTCTGGAAGCTGCAAGCTGGAGATGTCGAACCTGTAAAAA 1754
D 2307 GATGTTGATTATCAGACTCTTTAGAGGCATCTAAAGCTGGAGACTTTGGAACCTGTGAAGCAA 2366
QY 1755 CTGTGTAAGTGTTCAGAGCTCAACTGCAGACACATTTGAAGGGCTGCTGTACACCACTT 1814
D 2367 CTTTGCAGCTCTCAAAATGTGAATTTGTAGAGACTTAGAGGCCGCAATCCACGCCCTTA 2426
QY 1815 CATTTCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGTACAGCATGGAGCT 1874
D 2427 CACTTCGACAGGCTACAACCGCTGTCTGTGTAGAGTACCTGTGTACACCAAGGTGCC 2486
QY 1875 GATGTCATGCTAAAGATAAAGAGGCTTGTACCTTTGCACAATGCTGTCTTATGGA 1934
D 2487 GATGTCATGCCAAAGACAAGGGTGGCTTGGTCCCTTTCATATGCTGTTCATATGGA 2546
QY 1935 CATTATGAAGTTTCAGAACTTCTGTAAACATGGAGCAGTAGTTAATGTAGCTGATTTA 1994
D 2547 CACTATGAGTGTGCTGAGCTTTTAGTAAGCATGGGCTTCTGTCAATGTGGCGACTTA 2606
QY 1995 TGCATAATTTACACTTTTACATGAGCAGCAGCAAGAAAGAAATATGAAATTTGCAAACTT 2054
D 2607 TGAATATTTACCCCTCTCCATGAAGCAGCAGCTAAAGAAAGTATGAAATCTGCAAGCTC 2666
QY 2055 CTGCTCCAGCATGGTCAGACCCCTTACAAAAAAGCAGGATGGAATATCTCTTTGGAT 2114
D 2667 CTTTAAACATGAGCAGATCCAACTAAAAGAACAGAGATGGAATATACCTTTGGAT 2726
QY 2115 CTGTTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGAGATGCAGCTTTGCTA 2174
D 2727 TTGGTTAAAGAGAGGAGACACAGATATTCAGGACTTACTGAAAGGGATGCTGCTTTGTTG 2786
QY 2175 GATGCTGCCAAGAAAGGTTGTTTAGCCAGAGTGAAGAGTTGTCTCTCTGATATGTA 2234
D 2787 GATGCTGCCAAGAAAGGCTGCTGGCAGAGTGCAGAGCTGTACCCAGAGATATC 2846
QY 2235 AATTGCGCGATACCCAGGACACATTCACACCTTTACATTTAGCAGCTGTTTAAAT 2294
D 2847 AACTGCAGAGACACCCAGGCGAGAAATTCACCCCTCTGCACCTGCAGCAGCTATAAT 2906
QY 2295 AATTTAGAAGTTGCAGAGTATTTGTACACAGGAGCTGTATGTGAATGCCCAAGACAA 2354
D 2907 AACCTGGAAGTATGCTGAATATCTCTAGAGCATGGAGCTGATGTTAATGCCAGGACAAG 2366
QY 2355 GGAGGACTTATCTCTTACATAATGAGCATGTTAGGGCATGTAGATGTAGCAGCTCTA 2414
D 2967 GGTGTTTAAATCTCTTCAATGAGGCTCTTATGGCATGTTGACATAGCGGCTTTA 3026
QY 2415 CTAATAAGATATAGCATGTGTCAATGCCAGCAAAATGGGCTTTACAGCTTTGCAC 2474
D 3027 TTGATATAAATAACACACGCTGTGTAATGAACAGATAGTGGGCTTTACTCCCTCCAT 3086
QY 2475 GAAGCAGCCAAAAGGAGAACACAGCTTTGCTGCTTAGCCCATGGAGCTGAC 2534
D 3087 GAAGCAGCCCAAGAGGAGGAGCGAGCTGTGCGCCCTCTCTCTAGCGCATGGTGCAGAC 3146
QY 2535 CGACTCTTAAATATCAGGAGGACAAACACTTTAGATTTAGTTTTCAGCGCATGATGTC 2594
D 3147 CCCACCATGAAGAACAGGAGGCCAGCGCTCTGGATCTGGCAACAGCTGACGATATC 3206
QY 2595 AGCGCTCTTCTGACAGCAGCCATGCCCCCATGCTGTGCGCTCTTTGTACANGCTCAA 2654

Qy	492	GGGGCCCTTATTCTCTTATCAATGCAAGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC	551
Db	747	GGAGGCTCATCCGCTTCTAATAGCTCTTTTGGCCATGCTGAGTTGTGAGTCTG	806
Qy	552	CTTTTGGCAGATGCTGCAGACCCCAATGCTCAGATAATTGGAATTATATCTCCCTCCAT	611
Db	807	TTATTGTGCCAAGAGCTGATCCAAATGCCAGGATACTGGAATATACACTTCTGCAT	866
Qy	612	GAGCTGCAATTAAAGGAAAGATTGATTTGCATTTGCTTTTACAGCATFGAGCTGAG	671
Db	867	GAACTGCTATTAAAGGGAAGATCGATGTGCATTTGCTGCTGCAGCACGGAGCTGAC	926
Qy	672	CCAACCATCGAATACAGATGAAGACAGCAATTCGATTTAGCAGATCCATCTGCCAAA	731
Db	927	CCAACATCTGGAACTGATGGGAATCAGCCCTGGACCTGGCAGATPCTTTCAGAAAA	986
Qy	732	GCAGTGTACTGGTGAATATAAGAAAGATGAACCTCTAGAAAGTGCCAGGAGTGGCAAT	791
Db	987	GCTGCTTTACAGTGAATACAAAGAAAGCAACTCCTAGAACTGCTAGGAGTGGTAAT	1046
Qy	792	GAAGAAAAATGATGGCTTACTCACCATTAAATGTCACCTGCCACGCAAGTGATGGC	851
Db	1047	GAAGAAAAATGATGGCTTACTGACTCCTCTAAATGTGAATGCCATFGCAAGTGTGG	1106
Qy	852	AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG	911
Db	1107	CGAAAGTCGACTCCTTTACATCTAGCAGCGGCTACACAGAGTTCGNAATAGTTCAGCTT	1166
Qy	912	TTACTGCAACATGAGCTGATGTCATGCTAAAGATAAGGTGATGTGGTACCAATTACAC	971
Db	1167	CTTCTCAGCATGGTCTGATGTTCTATGCAAGAACAAAGGTGAGTGTGCTCTTTCTAT	1226
Qy	972	AATGCCCTGTTTATGTCCTATTGAAGTAACTGAACCTTTTGGTCAAGCATGGTGCTGT	1031
Db	1227	AATGCATGTTTCATATGGACATATGAAGTCA CAGAACTGCTACTAAAGCATGGAGCTGT	1286
Qy	1032	GTAATGCAATGAGCTVTGGCAATTCATCTCTTTCATGAGGAGGCTTCTAAGAACAGG	1091
Db	1287	GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGTCACGAGGCTGTCTCCAAGAACGT	1346
Qy	1092	GTTCAAGTATGTTCTCTCTTTAAGTTATGTGCGAGACCCAAACACTGCTCAATGTGCAC	1151
Db	1347	GTAGAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTAGTGTAGTCAACTGGCAT	1406
Qy	1152	AATAAAGTCTATAGACTTGGCTCCCACACACAGTTTAAAGAAAGATTAGCATATGAA	1211
Db	1407	GGCAAAAGTCTCTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTTTAA	1466
Qy	1212	TTTAAAGGCCACTCTGTTGTCGAAGCTGCACGAAGCTGATGTTATCTGNAATCAAAAA	1271
Db	1467	TTTAAAGGTCATTTCTTACTACAGCAGCGCAGAGAGCAGACTTAGCTAAAGTTAAAAA	1526
Qy	1272	CATCTCTCTCGGAAATGGTGAATTTCAAGCATCCTCAACACATGAAACAGCATTCAT	1331
Db	1527	ACACTGCTCTGGAATCATTAATTCAACAACCCGAGTCTCATGAACAGCACTGCAC	1586
Qy	1332	TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAATATGTGAACCTGTGCTAAGAAAA	1391
Db	1587	TGTGCTGTGGCTCTCTGCATCCCAACGTAACAAGTGAACAGATTTGTTACTTAGAAAA	1646
Qy	1392	GGAGCAAACTCAATGAAAGACTAAAGAAATTTCTTACTCTCTGACGTGGCATCTGAG	1451
Db	1647	GGAGCAATGTTAATGAAAAAATAAGATTTTCATGACTCTCCCTTGCATGTTGCAGCGAA	1706
Qy	1452	AAAGCTCATTAATGATGTTGTGAAGTGTGGTGAACATGAAGCAAAAGTTAATGCTCTG	1511
Db	1707	AGAGCCCATTAATGATGCTATGGAAGTCTTGCATAAAGCATGGCGCAAGATGATGCACTG	1766
Qy	1512	GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGCTGCTCATCTACAACTGC	1571
Db	1767	GACACCTCTGGTCAGACTGCTTTGCAATAGAGCGGCCCTAGCAGGCCACCTGCAGACCTGC	1826
Qy	1572	CGCTACTCTGAGCTATGGGTGTGATCCTAAACATATATCCCTTCAGGCTTTACTGCT	1631

Db	1827	CGCCTCCTGAGTTACGGCTCTGACCCCTCCATCATCTCTTACAAGGCTTCACAGCA	1886
Qy	1632	TTACAGATGGGAAATGAAATGTFACAGCAACTCCTCCAAGAGGTGTCTCATTAGGTAAAT	1691
Db	1887	GCACAGATGGCANTGACGAGTGCAGCAGATTCTGAGTGAGAGTACACCTATACGTACT	1946
Qy	1692	TCAGAGGCACAGACAAATTCCTGGAGCTGCAAGAGCTGGAGATGTCGAAACTGTAAAA	1751
Db	1947	TCTGATGTTGATATCGACTCTTAGAGGCAATCAAAGCTAGAGCTTGAAGCTTGGAAACTGTGAAG	2006
Qy	1752	AAACTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATCAAGGGGTGAGTCTACACCA	1811
Db	2007	CAACTTTGCAAGCTCTCAAAATGTAATGTAGAGACTTAGAGGGCGGCAATCCACGCC	2066
Qy	1812	CTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAAATATCTGCTACAGCATGGA	1871
Db	2067	TTACACTTCGCAGCAGGCTACAAACCGCTCTCTGTGTAGAGTACCTGCTACACCACGGT	2126
Qy	1872	GCTGATGTCGATGCTTAAAGATAAAGAGAGGCTTGTACCTTTGCAACAATGCATGTTCTTAT	1931
Db	2127	GGCGATGTCATGCCCAAGCAAGAGGTGGCTTGGTGCCCTTCATATGCGCTGTTTCATAT	2186
Qy	1932	GGACATATTGAAGTTCCGAAACTCTTGTTAAACATGGAGCAGTAGTAAATGTAGCTGAT	1991
Db	2187	GGACACTATGAGTGGCTGAGCTTTTAGTAAGCATGGGCTTCTGTCAATGTGGCGGAC	2246
Qy	1992	TTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAATTTGCAAA	2051
Db	2247	TTATGGAATTTTACCCTCTCCATGAAGCAGCAGCTTAAAGGAAAGTATGAATCTGCAG	2306
Qy	2052	CTTCTGCTCCAGCATGGTGCAGACCCCTACCACAAAAACAGGATGGAATACTCCCTTG	2111
Db	2307	CTCCTTTTAAACATGGAGCAGATCCAACTTAAAGACAGAGATGGAATACACCTTG	2366
Qy	2112	GATCTTGTGTAAGATGGAGATACAGATATTCGAAGTCTGCTTAGGGGAGATGCAGCTTG	2171
Db	2367	GATTTGGTAAAGGAAGGACACAGATATTCAGGACTTACTGAAAGGGATGCTGCTTG	2426
Qy	2172	CTAGATGTCGCAAGAAGGTTCTTTAGCCAGAGTGAAGAAGTTGTTCTCTCTGATAAT	2231
Db	2427	TTGGATGCTGCCAAGAGGCTGCCTGGCAAGAGTGCAGAGCTCTGTACCCAGAGAAAT	2486
Qy	2232	GTAATTTGCGCGATACCCAAAGCAGACATTCACACACCTTTTACATTTAGCAGCTGTTAT	2291
Db	2487	ATCAACTGCAGAGACACCCAGGGCAGAAATTCACCCCTCTGCACCTGGCAGCAGGCTAT	2546
Qy	2292	AATAATTTAGAAGTTCGACAGTATTTGTTACACACGGAGCTGATGTGAATGCCCAAGAC	2351
Db	2547	AATAAGCTGGAAGTAGCTGAATATCTTCTAGAGCATGGAGCTGATGTGAATGCCCAGGAC	2606
Qy	2352	AAAGGAGGACTTATTCCTTTACATATGCAGCATCTTACGG-----	2392
Db	2607	AAGGGTGGTTTAATTCCTCTTCAATATGCGGCATCTTATGGGGGCTGCCTGGCAAGAGTG	2666
Qy	2393	-----	2392
Db	2667	CAGAAGCTGTACCCAGAGAAATATCAACTGCAGAGACACCCAGGGCAGAAATTCAAAC	2726
Qy	2393	-----	2392
Db	2727	CCTCTCACCTGGCAGCAGGCTTAATAACCTGGAAGTAGCTGAATATCTCTTAGAGCAT	2786
Qy	2393	-----	2392
Db	2787	GGAGCTGATGTTAATGCCCGAGGACAAAGGTGGTTTAATTCCTTCATTAATGCGGCATCT	2846
Qy	2393	-----	2392
Db	2847	TATGGTAGTAAAGTTGGATTCCAGAGCCTCTTCCAGCTTGTGTGAATGATTAATG	2906
Qy	2393	-----GCATGT-----	2398

Db 2907 AGACCATGCATGTGGAATTCGATTAACATAATGTAAGGCATTATATAAAATGCAAGCATGT 2966
QY 2399 AGATGTACAGAGCTTACTAATAAAGTATAATGCGATGTGTCATGCGCAGCAGCAATGGGC 2458
Db 2967 TGACATAGCGGCTTTATTTGATAAAATACACACCGTGTGTAATGCAACAGATAAGTGGC 3026
QY 2459 TTTCACACTTTTGACGAAGCAGCCAAAGAGGAGCAGCAACACAGCTTTGTGCTTTGTGCT 2518
Db 3027 GTTTACTCCCTCCATGAAGCAGCCGAGAAAGGAGGAGCAGCTGTGCGCCCTCCTCCT 3086
QY 2519 AGCCCATGAGGTGACCGAGCTCTTTAAATTCAGGAAGGAGCAACACCTTTAGATTTAGT 2578
Db 3087 AGCGCATGCTGAGAGCCCGCAGCATGAAGAACAGGAGGAGCCAGCGCTCTGGATCGGC 3146
QY 2579 TTCAGGGATGATGTCAGGCTCTCTCT----- 2605
Db 3147 AACAGCTGACGATATCAGAGCTTTGCGATGTTGACATAGCGGCTTTATTTGATAAAATACAA 3206
QY 2606 ----- 2605
Db 3207 CACGTGTGTAATGCAACAGATAAGTGGGCGTTTACTCCCTCCATGAAGCAGCCAGAA 3266
QY 2606 -----GACAGCAGCATGCCCGCCTCTGCTGCGCCTCTTTGTTACAGCCCTCAAG----- 2655
Db 3267 AGGAAGGAGCGAGCTGTGCGCCCTCTCTAGCGCATGGTGCAGACCCCGCAGCATGAAGAA 3326
QY 2656 -----TGCTCAA 2662
Db 3327 CCAGGAAGGCCAGAGCCCTCTGGATCTGGCAACAGCTGACGATATCAGAGCTTTGCTGAT 3386
QY 2663 TGGTGTGGAAGCCCGAGGAGCCATGCG----- 2689
Db 3387 AGATGCCATGCCCCCAGAGGCGCTTACCTACCTGTTTAAACCTCAGGCTACTGTAGTGAG 3446
QY 2690 AGATGCTCTCTCTTCAGGTCCATCTAGGCCATCAAGCCCTTTCTGCGAGCCAGCAGCTTTGA 2749
Db 3447 TGCCCTCTCTGATCTCACCAGCATCACCCCTCTCTGCGCTCGGCTGCCAGCAGCATAGA 3506
QY 2750 CAACCTATCTGGAGTTTTCAGAACTCTCTCAGTAGTTAGTTCAAGTGAACAGAGGG 2809
Db 3507 CAACCTCAGTGGCTTTAGCAGAGTTGGCGTAGGAGGCTCCATGAGCGGGATGG 3566
QY 2810 TGCTTCCAGTTTGGAGAAAGAGG-----GTTCCAGGAGTAGATTTTAGCATAACTCA 2863
Db 3567 CGCCGCGGAACAGAAAGGAGGAGAGAGTTGCTGCTTGACATCAATATCAGCCA 3626
QY 2864 ATTCGTAAGGAATCTTGGACCTTGACCACTTAATGATATATTTGAGAGAGACAGATCAC 2923
Db 3627 ATTTCTAAAAGCCCTTGGCCCTTGAACACCTTCGGGATATCTTTGAAACAGAACAGATTAC 3686
QY 2924 TTTGGATGTATTAGTTGAGATGGGCACAAAGGAGCTGAAGGAGATTGGAATCAATGCTTA 2983
Db 3687 ACTAGATGTTTGGCTGATATGGGTCTGATGAGAGTTGAAAGAAATAGGCATCAATGCAATA 3746
QY 2984 TGGACATAGGCACAACTAATTTAAAGGAGTCGAGAGCTTATCTCCGAGCAACAAGGTCT 3043
Db 3747 TGGGCACCCCAAAATTAATCAAGGAGCTAGAAAGACTCTTTAGGTGGCAACAAGGCAC 3806
QY 3044 TAACCATATATTTAACTTTGAACACCTCTGGTAGTGGCAATCTTTATAGATCTGCTGCC 3103
Db 3807 CAATCCTTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGTGCTTTGCTCC 3866
QY 3104 TGATGATAAGAGTTTTCAGCTGTGGAGGAAGAGATGCAAGTACAGTTTCGAGAGCACAG 3163
Db 3867 AGAAGATAAAGAAATATCACTCAGTGGAGAAAGAGATGCAAGTACTATTCGAGAACACAG 3926
QY 3164 AGATGAGGTCTATGCAAGTGGAACTTTCAACAGATACAAATTTCTCAAGATTCAGAAGGT 3223
Db 3927 AGATGCTGTTATGCTGGCGGCATCTTCAACAGATACAAATGCTATTCGAATTTCAAAAAGT 3986
QY 3224 TTGTACAGAAACTATGGGAAGATACACTCACCGGAGAAAGAGTTTCTGAAGAAA 3283
Db 3987 TGTCAACAAGAGTTTGAGGGAGCGGTTCTGCCACCGCAGAGAAGGAGTGTCTGTGAGAGAA 4046

QY 3284 CCACAACCATGCCAATGCAAGCAATGCTATTTCATGGTCTCCTTTTGTGAATGCAATATAT 3343
Db 4047 TCACAACCATCACATGACGCGCATGTGTTTCATGGTTCCTTCCTTATTAATGCCATATAT 4106
QY 3344 CCACAAGGCTTTTGATGAAAGGCGATCGCTACATAGGTGATGTTTGGAGCTGGCATTTA 3403
Db 4107 TCATAAAGGCTTTTGATGAGCGACATGATACATAGGAGGAATCTTTGGGCGCGGATTTA 4166
QY 3404 TTTTGTGAAAACCTTCCAAAAGCAATCAATATGATATGTAATGGAATGGAGGAGGTACTGG 3463
Db 4167 TTTTGTGAAAACCTTCCAAAAGCAATCAATATGTTATGTAATGGAGGAGGAACAGG 4226
QY 3464 GTGTCCAGTTTCACAAGACAGATCTTGTTCATTTGCTCCACAGCAGCTGCTCTTTTGGCG 3523
Db 4227 CTGCCCTACACACAGCAGAGTCATGCTATATATGTCACAGACAAATGCTCTTCTGTAG 4286
QY 3524 GGTAAACCTTGGGAAGTCTTTTCTGCAAGTTCAGTGCATGCAATGAAATGGCAGATTTCTCTCC 3583
Db 4287 AGTGACCCCTTTGGGAAATCTTCTGCAAGTTAGCACCATGAAATGGCCACGCGCTCC 4346
QY 3584 AGTCACTCACTCACTCACTGCTAGGCCAGCTGTAATGGCTAGCATTTAGCTCAATATGT 3643
Db 4347 AGGCACCACTCACTCACTGCTAGGCCAGCGGCTCAATGGCTGGCATATGCTGAATATGT 4406
QY 3644 TATTACAGAGGAGCAAGGCTTATCTCTGAGTATTTAAATTAATTTACCAATATGAGGCC 3703
Db 4407 CATCTACAGAGGAGCAAGCAGCATACCAGAGTATCTTATCACTTACCAGATCATGAAGCC 4466
QY 3704 TGAAG 3708
Db 4467 AGAAG 4471

RESULT 5

US-09-964-899-40
; Sequence 40, Application US/09964899
; Patent No. US20020174446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-964-899-40

Query Match 31.3%; Score 1192.6; DB 9; Length 2409;

Best Local Similarity 67.4%; Pred. No. 9.9e-314; Indels 1128; Gaps 4;
Matches 2372; Conservative 0; Mismatches 19;

QY 225 ATGTGCGGTCCGCTGCGCGCGGGGAGCGGCTTGGCGAGCGCGCGCGGCGGCGGCGG 284
Db 1 ATGTGCGGTCCGCTGCGCGCGGGGAGCGGCTTGGCGAGCGCGCGCGGCGGCGGCGG 60
QY 285 GTGAGCGCGCGCGCGCGAGAGCTGTTGAGCGGTGCGCGCAACGGGGAGCTGGGAACGAGTC 344
Db 61 GTGAGCGCGCGCGCGCGAGAGCTGTTGAGCGGTGCGCGCAACGGGGAGCTGGGAACGAGTC 120
QY 345 AAGAGGCTGGTGACGCTGAGAAGTGAAACAGCGCGACGCGCGGCGGCAAGAAATCCACC 404
Db 121 AAGAGGCTGGTGACGCTGAGAAGTGAAACAGCGCGGCGGCGGCGGCGGCAAGAAATCCACC 180

QY 405 CGCTGCACTTCGCCCGCAGGTTTGGGCGGAAGACGTAGTTGAATATTGTCCTCAGAAAT 464
Db 181 CGCGTCACACTTCGCCCGCAGGTTTGGGCGGAAGACGTAGTTGAATATTGTCCTCAGAAAT 240
QY 465 GGTGCAAAATGTCACACGACGTGATGGGCGCTTATTCCTCTTCATATATGCATGCTCT 524
Db 241 GGTGCAAAATGTCACACGACGTGATGGGCGCTTATTCCTCTTCATATATGCATGCTCT 300
QY 525 TTTGGTCACTGCAAGTAGTCAATCTCCTTTTGGCAGACATGGTGCACACCCCAATGCTCA 584
Db 301 TTTGGTCACTGCAAGTAGTCAATCTCCTTTTGGCAGACATGGTGCACACCCCAATGCTCA 360
QY 585 GATAATTGAATATATCTCCTCCATGAAGTGCATTAAGAGGAAGATGATGTTTGC 644
Db 361 GATAATTGAATATATCTCCTCCATGAAGTGCATTAAGAGGAAGATGATGTTTGC 420
QY 645 ATTGTGCTCTTACAGCATGAGCTGAGCAACCATCCGAAATACAGATGAAGGACAGA 704
Db 421 ATTGTGCTCTTACAGCATGAGCTGAGCAACCATCCGAAATACAGATGAAGGACAGA 480
QY 705 TTTGGATTTAGCAGATCCATCTGCCAAGCAGTCTTACTTGGTAAATATATATATATAT 752
Db 481 TTTGGATTTAGCAGATCCATCTGCCAAGCAGTCTTACTTGGTAAATATATATATATATAT 540
QY 753 -----AAGAAAGATGAATCTTAGAAGTGCAGAGTGGCAATGAAGAAAATGATG 806
Db 541 TATTCAGGAAGCCCTGTAAGAAACAACTTGCAGGAGTGGCAATGAAGAAAATGATG 600
QY 807 GCTCTACTCACACCATTAATGCTCAACTGCCACGCAAGTGTGGCAGAAAGTCAACTCA 866
Db 601 GCTCTACTCACACCATTAATGCTCAACTGCCACGCAAGTGTGGCAGAAAGTGTGGCAG 648
QY 867 TTACATTTGGCAGCAGGATATACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGA 926
Db 649 ----- 648
QY 927 GCTGATGTCATGCTTAAGATTAAGGTGATCTGGTACCATTACACAAATGCCCTGTTCTAT 986
Db 649 ----- 648
QY 987 GGTCAATTATGAAGTAACATGAACCTTTTGGTCAAGCATGGTCCCTGTGTAATGCAATGGAC 1046
Db 649 -----AAGCATGGTCCCTGTGTAATGCAATGGAC 678
QY 1047 TTGTGGCAATTCACCTCTTCATGAGGAGCTTCTAAGAACAGGGTGAAGTATGTTCT 1106
Db 679 TTGTGGCAATTCACCTCTTCATGAGGAGCTTCTAAGAACAGGGTGAAGTATGTTCT 738
QY 1107 CTCTCTTAAGTTATGGTGCAGACCCCAACACTGCTCAATTTGTCACAAATAAAGTCTATA 1166
Db 739 CTCTCTTAAGTTATGGTGCAGACCCCAACACTGCTCAATTTGTCACAAATAAAGTCTATA 798
QY 1167 GACTTGGCTCCACACACAGTTAAAGAAAGATTAGCATATGAATTTAAGGCCACTCG 1226
Db 799 GACTTGGCTCCACACACAGTTAAAGAAAGATTAGCATATGAATTTAAGGCCACTCG 858
QY 1227 TTGCTGCAAGTGCACGAGACGTGATGTTACTCGAATCAAAAACATCTCTCTCTGAA 1286
Db 859 TTGCTGCAAGTGCACGAGACGTGATGTTACTCGAATCAAAAACATCTCTCTCTGAA 918
QY 1287 ATGTTGAATTTCAAGCATCTCAACACATGAACAGCATTTGCTGCTGCTCATCT 1346
Db 919 ATGTTGAATTTCAAGCATCTCAACACATGAACAGCATTTGCTGCTGCTCATCT 960
QY 1347 CCATATCCAAAAGAAAGCAATATGTGAACCTGTTGCTTAAGAAAGGACAAACATCAAT 1406
Db 961 ----- 960
QY 1407 GAAAGACTAAGAAATCTTGACTCTCTGACAGTGGCATCTGAGAAAGCTCATATGAT 1466
Db 961 ----- 960
QY 1467 GTTGTGAAGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTGGATAATCTTGGTCAG 1526

Db 961 ----- 960
QY 1527 ACTTCTCTACACAGAGCTGCATATATGTTGGTCACTCTACAAACCTGCCCTACTCTCAGC 1586
Db 961 ----- 960
QY 1587 TATGGTGTGATCCTAAAT 1646
Db 961 ----- 960
QY 1647 GAAAATGTACAGCAACTCTCTCAAGAGGATATCTATATAGTAAATTCAGAGGACAGACA 1706
Db 961 ----- 960
QY 1707 CAATTGCTGGAAGCTGCAAGGCTGGAGATGTGCAAACTGTAAAAAACTGTGTACTGTT 1766
Db 961 -----AAACTGTGTACTGTT 975
QY 1767 CAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCACTCTACACACTTTCATTTTGCAGCT 1826
Db 976 CAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCACTCTACACACTTTCATTTTGCAGCT 1035
QY 1827 GGTATACAGAGTGTCCGTGTGGAAATATCTCTACAGCATGGAGCTGATGTGCATGCT 1886
Db 1036 GGTATACAGAGTGTCCGTGTGGAAATATCTCTACAGCATGGAGCTGATGTGCATGCT 1095
QY 1887 AAAGATAAGGAGGCGCTTGTACCTTTGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1946
Db 1096 AAAGATAAGGAGGCGCTTGTACCTTTGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1155
QY 1947 GCAGAACTCTTTGTTAAACATGGAGCAGTAAATATGAGCTGATTTATGGAATTTTACA 2006
Db 1156 GCAGAACTCTTTGTTAAACATGGAGCAGTAAATATGAGCTGATTTATGGAATTTTACA 1215
QY 2007 CCTTTACATGAAGCAGCAGCAAAAGAAATATGAAATTTGCAAACTTCTGCTCCAGCAT 2066
Db 1216 CCTTTACATGAAGCAGCAGCAAAAGAAATATGAAATTTGCAAACTTCTGCTCCAGCAT 1275
QY 2067 GGTGCAGACCTTACCACCAAAACAGGATGCAATACCTTTCGATCTTCTTAAAGAT 2126
Db 1276 GGTGCAGACCTTACCACCAAAACAGGATGCAATACCTTTCGATCTTCTTAAAGAT 1335
QY 2127 GGAGATACAGATATTTCAAGATCTGCTTAGGGGAGATGAGCTTTGCTAGATGCTGCCAAG 2186
Db 1336 GGAGATACAGATATTTCAAGATCTGCTTAGGGGAGATGAGCTTTGCTAGATGCTGCCAAG 1395
QY 2187 AAGGTTGTTTAGCCAGAGTGAAGAGTTGCTTCTCTGATTAATGTAATTTGCCGCGAT 2246
Db 1396 AAGGTTGTTTAGCCAGAGTGAAGAGTTGCTTCTCTGATTAATGTAATTTGCCGCGAT 1455
QY 2247 ACCCAAGGCAGACATTTCAACACTTTTACATTTAGCAGCTGGTTATATAATTTAGAAGTT 2306
Db 1456 ACCCAAGGCAGACATTTCAACACTTTTACATTTAGCAGCTGGTTATATAATTTAGAAGTT 1515
QY 2307 GCAGAGTATTTGTTTACAACAGGAGCTGATGTAATGCCCAAGACAAAGAGGAGCTTAT 2366
Db 1516 GCAGAGTATTTGTTTACAACAGGAGCTGATGTAATGCCCAAGACAAAGAGGAGCTTAT 1575
QY 2367 CCTTTACATATGACAGCATTTACGGGCTGATGATGTAGCAGCTCTACTAATAAGTAT 2426
Db 1576 CCTTTACATATGACAGCATTTACGGGCTGATGATGTAGCAGCTCTACTAATAAGTAT 1601
QY 2427 AATGATGTGTCAATGCCACGACAAATGGGCTTTTACACACTTTTGCACGAGCAGCCCAA 2486
Db 1602 ----- 1601
QY 2487 AAGGACGACACAGCTTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAA 2546
Db 1602 ----- 1601
QY 2547 AATCAGGAAGGACAAACACTTTTAGATTTAGTTTTCAGCGGATGATGTAGCGCTCTTCTG 2606

Db 1602 ----- 1601
Qy 2607 ACAGAGCCATGCCCCCATCTGCTCTGCCCTCTTTGTACAAAGCTCAAGTGCTCAATGGT 2666
Db 1602 ----- 1601
Qy 2667 GTGAGAGCCCGAGGCGCACTGCAGATGCTCTCTTTCAGGTCCATCTAGCCCATCAAGC 2726
Db 1602 ----- 1601
Qy 2727 CTTTCTGCGCCAGCAGCTTTGACAACTTATCTGGAGTGTTCAGAACTGTCTTCAGTA 2786
Db 1602 ----- 1601
Qy 2787 GTTAGTTCAAGTGGACAGAGGGTCTCCAGTTTGGAGAAAAGGAGGTTCAGGAGTA 2846
Db 1602 ----- 1601
Qy 2847 GATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAAATGGATATATT 2906
Db 1602 ----- 1601
Qy 2907 GAGAGAGAACAGATCACTTTTGGATGTATTAGTTGAGATGGGCGCACAAAGAGCTGAAGGAG 2966
Db 1602 -----GATCACCTTGGATGATTAGTTGAGATGGGCGCACAAAGAGCTGAAGGAG 1650
Qy 2967 ATTGGAATCAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGACTTATC 3026
Db 1651 ATTGGAATCAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGACTTATC 1710
Qy 3027 TCCGCACAAAGGCTTAAACCCATATTAACTTTGAACACCTCTGTGAGTGGAAACAATT 3086
Db 1711 TCCGCACAAAGGCTTAAACCCATATTAACTTTGAACACCTCTGTGAGTGGAAACAATT 1770
Qy 3087 CTTATAGATCTCTCTCTGATGATAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAGT 3146
Db 1771 CTTATAGATCTCTCTCTGATGATAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAGT 1830
Qy 3147 ACAGTTCGAGACACAGAGATGGAGTCTATGAGTGGAACTTCAACAGATACAAATATT 3206
Db 1831 ACAGTTCGAGACACAGAGATGGAGTCTATGAGTGGAACTTCAACAGATACAAATATT 1890
Qy 3207 CTCAGATTCAGAAGGTTTGTAAACAGAACTATGGGAAGATACACTCACCGGAGAAAA 3266
Db 1891 CTCAGATTCAGAAGGTTTGTAAACAGAACTATGGGAAGATACACTCACCGGAGAAAA 1950
Qy 3267 GAAGTTTCTGAAGAAACACACAACTATGCCAATGAACGAATGCTATTTCATGGGTCTCT 3326
Db 1951 GAAGTTTCTGAAGAAACACACAACTATGCCAATGAACGAATGCTATTTCATGGGTCTCT 2010
Qy 3327 TTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAGGCATGCTACATAGTGGTATG 3386
Db 2011 TTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAGGCATGCTACATAGTGGTATG 2070
Qy 3387 TTTGGAGCTGGCATTTATTTTGTGAAAACTCTTCCAAAGCAATCAATATATATGGA 3446
Db 2071 TTTGGAGCTGGCATTTATTTTGTGAAAACTCTTCCAAAGCAATCAATATATATGGA 2130
Qy 3447 ATTGGAGAGGACTGGGTGTCAGTTTCACAAAGACAGATCTGTACATTTGCCACAGG 3506
Db 2131 ATTGGAGAGGACTGGGTGTCAGTTTCACAAAGACAGATCTGTACATTTGCCACAGG 2190
Qy 3507 CAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGCTCTTCTGCAGTTCAGTGCATGAAA 3566
Db 2191 CAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGCTCTTCTGCAGTTCAGTGCATGAAA 2250
Qy 3567 ATGGCACAATCTCTCCAGGTCATCACTCAGTCACTGTGGGCCAGGTGTAATGGCCTA 3626
Db 2251 ATGGCACAATCTCTCCAGGTCATCACTCAGTCACTGTGGGCCAGGTGTAATGGCCTA 2310
Qy 3627 GCATTAGCTGAATATGTTTATTTACAGAGGAGACAGGCTTATCCTCAGTATTTAATTACT 3686
Db 2311 GCATTAGCTGAATATGTTTATTTACAGAGGAGACAGGCTTATCCTCAGTATTTAATTACT 2370

Qy 3687 TACCAGATTATGAGGCTGAAGGTATGGTCGATGGATAA 3725
Db 2371 TACCAGATTATGAGGCTGAAGGTATGGTCGATGGATAA 2409

RESULT 6

US-09-833-381-841
; Sequence 841, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 841
; LENGTH: 1069
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(1069)
; OTHER INFORMATION: n = A,T,C or G
; US-09-833-381-841

Query Match 12.3%; Score 467.6; DB 10; Length 1069;

Best Local Similarity 98.8%; Pred. No. 8.1e-117;

Matches 492; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy 1662 CTCTCCACAGAGGTATCTCATTTAGGTAAATTCAGAGCGACAGACAAATTCGTGGAAGCT 1721
Db 567 CCCTCTAGAGGTATCTCATTTAGGTAAATTCAGAGCGACAGACAAATTCGTGGAAGCT 626
Qy 1722 GCAAAGCTGGAGATGCGAAACTGTAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 1781
Db 627 GCAAAGCTGGAGATGCGAAACTGTAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 686
Qy 1782 AGAGACATTTGAAGGGCTCAGTCTACACCACCTTCATTTTCGACGTGGGTATACAGAGTG 1841
Db 687 AGAGACATTTGAAGGGCTCAGTCTACACCACCTTCATTTTCGACGTGGGTATACAGAGTG 746
Qy 1842 TCCGTGTGGAATATCTGCTACAGCATGGAGCTGATGCTGCTAAAGATAAAGGAGGC 1901
Db 747 TCCGTGTGGAATATCTGCTACAGCATGGAGCTGATGCTGCTAAAGATAAAGGAGGC 806
Qy 1902 CTTGTACCTTTGCAATGCTGCTTATGGACATTTGAAGTTGCAGAACTTCTTTGTT 1961
Db 807 CTTGTACCTTTGCAATGCTGCTTATGGACATTTGAAGTTGCAGAACTTCTTTGTT 866
Qy 1962 AAACATGAGCAGTAGTTAATGTAGCTGATTTAGGAAATTTACACCTTTACATGAAGCA 2021
Db 867 AAACATGAGCAGTAGTTAATGTAGCTGATTTAGGAAATTTACACCTTTACATGAAGCA 926
Qy 2022 GCAGCAAAAGCAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACC 2081
Db 927 GCAGCAAAAGCAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACA 986
Qy 2082 AAAAAAACA -GGATGGAATACTCCCTTTGGATCTTTGTTTAAAGATGAGATACAGATAT 2140
Db 987 AAAAAAACAAGGATGGAATACTCCCTTTGGATCTTTGTTTAAAGATGAGAT -CAGATAT 1045
Qy 2141 TCAAGATCTCTAGGGG 2158
Db 1046 TCAAGATCTCTAGGGG 1063

RESULT 7

US-09-833-381-1153

; Sequence 1153, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1153
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1153

Query Match 12.1%; Score 461; DB 10; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.8e-115;
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 572 CCCCAATGCTCGAGATAATTTGGAAATTAATCTCTCCATGCAAGCTGCAATTAAGGAAA 631
Db 5 CCCCAATGCTCGAGATAATTTGGAAATTAATCTCTCCATGCAAGCTGCAATTAAGGAAA 64
QY 632 GATTGATGTTTGGCATTGCTGTGTACAGCATGGAGCTGAGCCAAACCATCCGAAATACAGA 691
Db 65 GATTGATGTTTGGCATTGCTGTGTACAGCATGGAGCTGAGCCAAACCATCCGAAATACAGA 124
QY 692 TGGAGGACAGCATGTTGATTTACAGATCCATCTGCCAAAGCAGTGTCTTACTGGTGAATA 751
Db 125 TGGAGGACAGCATGTTGATTTACAGATCCATCTGCCAAAGCAGTGTCTTACTGGTGAATA 184
QY 752 TAAGAAAGATGAATCTCTAGAAAGTGCAGGAGTGGCAATGAAGAAATGATGGCTCT 811
Db 185 TAAGAAAGATGAATCTCTAGAAAGTGCAGGAGTGGCAATGAAGAAATGATGGCTCT 244
QY 812 ACTCACACCATTAATGTCCTCAACTGCCAGCAGTGTGGCAGAAAGTCAACTCCATTACA 871
Db 245 ACTCACACCATTAATGTCCTCAACTGCCAGCAGTGTGGCAGAAAGTCAACTCCATTACA 304
QY 872 TTTGGCAGCAGGATATAACAGATAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGA 931
Db 305 TTTGGCAGCAGGATATAACAGATAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGA 364
QY 932 TGTCATGCTTAAAGATAAAGGTGATCTGTGTACCATACCAATGCCCTGTTCTTATGGTCA 991
Db 365 TGTCATGCTTAAAGATAAAGGTGATCTGTGTACCATACCAATGCCCTGTTCTTATGGTCA 424
QY 992 TTATGAAGTAAGTGAATCTTTGGTCAAGCATGGTGCCTGTG 1032
Db 425 TTATGAAGTAAGTGAATCTTTGGTCAAGCATGGTGCCTGTG 465

RESULT 8

US-10-040-739-1179
; Sequence 1179, Application US/10040739
; Patent No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John
; Lavallie, Edward
; Racie, Lisa
; Merberg, David
; Treacy, Maurice
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
; NUMBER OF SEQUENCES: 1519
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge

STATE: Massachusetts
COUNTRY: U.S.A
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/040,739
FILING DATE: 07-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/036,520
FILING DATE: 03-JUN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 1179:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 1179:
US-10-040-739-1179
Query Match 8.5%; Score 323; DB 9; Length 353;
Best Local Similarity 99.7%; Pred. No. 7.7e-78;
Matches 323; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3252 ACTCACCGGAGAAAGATTTCTTGAGAAACACCAACCATGCCAATGAACGAATGCTA 3311
Db 21 ACTCACCGGAGAAAGATTTCTTGAGAAACACCAACCATGCCAATGAACGAATGCTA 80
QY 3312 TTTTCATGGGTCTCTTTTGTGAATGCAATATCCAAAGGCTTTGTAGAAAGCATGCG 3371
Db 81 TTTTCATGGGTCTCTTTTGTGAATGCAATATCCAAAGGCTTTGTAGAAAGCATGCG 140
QY 3372 TACATAGGTGATGTTGAGCTGGCATTATTTGCTGAAACCTCTCCAAAGCAAT 3431
Db 141 TACATAGGTGATGTTGAGCTGGCATTATTTGCTGAAACCTCTCCAAAGCAAT 200
QY 3432 CAATATGATATGGAATTTGAGGAGGTACTGGGTGCCAGTTCCAGTTCCAAAGACAGATCTGT 3491
Db 201 CAATATGATATGGAATTTGAGGAGGTACTGGGTGCCAGTTCCAGTTCCAAAGACAGATCTGT 260
QY 3492 TACATTTGCCACAGGCGCTGCTCTTTTGCCTGGGTAACTTGGGAAAGTCTTTCCCTGCAG 3551
Db 261 TACATTTGCCACAGGCGCTGCTCTTTTGCCTGGGTAACTTGGGAAAGTCTTTCCCTGCAG 320
QY 3552 TTCAGTGCATGAAATGGCAGAT 3575
Db 321 TTCAGTGCATGAAATGGCAGAT 344
RESULT 9
US-09-841-835-12
; Sequence 12, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRP1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor

```

CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841.835
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/196,387
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-09-841-835-12

```

	Query Match	3.28;	Score 121.2;	DB 10;	Length 523;
	Best Local Similarity	74.3%;	Pred. No.	7.7e-23;	
	Matches 153;	Conservative 0;	Mismatches 53;	Indels 0;	Gaps 0;
Qy	2190	GGTGTGTTAGCCAGAGTGGAAGAAGTTGCTTCCTCGATATAATGTAAATTGCCGAGTACC	2249		
Db	1	GGCTGCCTGGCAAGAGTSCAAGAAGCTCTGTACCCCAGAGAATATCAACTGCAGAGACACC	60		
Qy	2250	CAAGCAGACAYTTCACACACTTTACATTTACAGAGCTGGTTATATAATTTAGAAGTTGCA	2309		
Db	61	CAGGSCAGAAAATTC AACCCCTCTGCACCTGGCAGCAGGCTATAATAAGCTGGAAGTAGCT	120		
Qy	2310	GAGTATTTGTTACAAACACGGAGCTCATCTGTAATGCCCAAGACAAAGGAGGACTTTATTCT	2369		
Db	121	GAATATCTCTAGACCATGGAGCTGATGTTAATGCCCAGGACACAGGGTGTTTAATTCCT	180		
Qy	2370	TTACATAATCAGCATCTTTACGGGCA	2395		
Db	181	CTTCATAATGGCGACTCTTATGGGTA	206		

```

RESULT 10
US-09-783-590-10862
; Sequence 10862, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: FO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21

```

```

; NUMBER OF SEQ ID NOS: 12495
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 10862
; LENGTH: 304
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (23)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (36)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (74)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (124)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (133)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (147)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (148)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (238)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (241)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (242)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (248)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (270)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (298)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (299)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-10862

```

	Query Match	3.1%	Score 117.6;	DB 10;	Length 304;
	Best Local Similarity	82.3%;	Prod. No. 5e-22;		
	Matches 237;	Conservative	0;	Mismatches 36;	Indels 15; Gaps 9;
Qy	3179	AGGTGGAAATCTTCAACACATACAATATTCTCAAGATTTCAGAAAGGTTTCTGAACAAGAAACT	3238		
Db	6	AGGTGGAAATCTTCAACANATACAATATTCTNAAGATTCAAAAGTTTCTGAACAAGAAACT	65		
Qy	3239	ATGGGAAAG-ATACACTCACC GGAGAAAGATTCTTG-AGAAAACCAACCACTGCC	3296		
Db	66	ATGGGAAANAATACACTCACC GGAGAAAGATTCTTGAAGAAGAAACCAACCACTGNC	125		
Qy	3297	--AATGAACGAATGCTATTTCAT---GGGTCTCCTTTTGTGAATGCAATTAT-CCACAA	3349		
Db	126	CAATGAANC GATTGCTAATTTNNATGGGTCTCCCTTTTGTGNAATGCAATTATGCCACA	185		
Qy	3350	AGGCCTTTCA-TCAAAGGCATCGT-ACATAGTGGTATGTTTGGAGCTGGCAATTTATTTT	3407		
Db	186	AGGCCTTTGATTCAAAAGCATCGTAAACATAGTGGTATGTTTGGGCCTGGGCCTTTNNTTT	245		

; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,888
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01329
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,905
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01354
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,891
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01339
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01340
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,874
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01334
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,898
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01320
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,853
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01349
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,902
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01239
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,870
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01348
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,882
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01347
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,896
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01307
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,864
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01341
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,856
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01336
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,868
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01312
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 12
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE
; LOCATION: (13)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (14)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1287)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-908-711-12

Query Match 2.0%; Score 77; DB 10; Length 1299;
Best Local Similarity 45.5%; Pred. No. 1.6e-10;
Matches 337; Conservative 4; Mismatches 394; Indels 2;

Qy 1384 TAAGAAAGGAGCAACATCAATGAAAGAGCTAAGAAATCTTGACTCCTCTGCACGTGG 1443
Db 203 TAAAAATGGAGTTGACATCAACATTTGCAATCAGATGGTTGAGCGCTCTCCACCTTG 262
Qy 1444 CATCTGAGAAAGCTCATAATGATGTTGAAAGTAGTGGTGAACATGAACAGCAAGGTGA 1503
Db 263 CTTCCAAAGAGGCCATGTAGAGGTTGTTCTGAGCTGCTGCAGAGAGAAAGCAATGTGG 322
Qy 1504 ATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTTCATCTAC 1563
Db 323 ATGCAGCTTACAAAGAAAGAAACACAGCAITTCACATCGCATCTTTGGCTGGGCAAGCAG 382
Qy 1564 AAACCTGCCGCTACTCCTGAGCTATGGGTGATGCTTCAACATTTATTCCTTTCAGGGCT 1623
Db 383 AGGTGGTAAAGTCTTGGTTACAAATGGAGCCAAATGCAATGCACATCTCAGAAATGGTT 442
Qy 1624 TTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCTCAAGAGGGTATCT--- 1680
Db 443 TCAGCCCATTTGATATGTCARCCAGGAAATCACCTGGAAGTTGTGAAGTTTCTCTTG 502
Qy 1681 --CATTAGGTAAATTCAGAGGAGCAGACACAATTCGTGGAAGCTCAAGAGCTGGAGATGT 1738
Db 503 ACAATGGTCAAGCCAGARCCCTARCCACAGARGATGGCTTCACACCAATTCGACGTGGCTT 562
Qy 1739 CGAAACTGTAAAAAAGTGTGACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGCG 1798
Db 563 TCAACAAGGTCACGACCAAGTCGTTTCGCTCCTGTAGAGAAATGCACCAAGGAAAG 622
Qy 1799 TCAGTCTA-CACCACTTCATTTTGCAGCTGGGTATTAACAGAGTGTCCGTTGGTGAATATC 1857
Db 623 TGCCTCTCCAGCTCTTCATATCGCGCCGAAAGACGACACGAAAGCCGCCCTGC 682
Qy 1858 TGCTACAGCATGGAGTGTGATGTCATGCTTAAAGATAAAGAGGCGCTTGTACCTTTGCACA 1917
Db 683 TGCTGAGAAATGACAAATGAGATGGAATGGAATCAAGAGTGGCTTCACCTCCGCTCCACA 742
Qy 1918 ATGCATGTTCTTATGACATTTATGAAGTTGCAGAACTTCTTTTAAACATGGAGCAGTAG 1977
Db 743 TAGCTCTCCTATGGAATATCAATGTAGCCACGTTGCTGTAAACGARGCGCTGTG 802
Qy 1978 TTAATGTAGCTGATTTATGGAATTTACACCTTTACATGAACGACGACGAAAGAAAT 2037
Db 803 TGGATTTCCCGCAAGGAATGACATCACCTTTTACATTTGATGATCAAAAGAGGAATG 862
Qy 2038 ATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAAGAGGATG 2097
Db 863 CAATATGTTAAACTATTTGCTGCTGATCGAGGAGCTAAATTCGATGCCAAACAGGATG 922
Qy 2098 GAAATACCTTTTGGATCTTG 2118
Db 923 GTCTGACACCACTGCCTCTG 943

RESULT 14
US-09-947-199-3
; Sequence 3, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:

APPLICANT: Raju, Jeyaseelan
TITLE OF INVENTION: NOVEL CARX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: MNI-068CP2
CURRENT APPLICATION NUMBER: US/09/947,199
PRIOR FILING DATE: 2001-09-05
PRIOR APPLICATION NUMBER: 60/111,938
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/291,839
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: 09/458,457
PRIOR FILING DATE: 1999-12-10
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 2505
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2505)
US-09-947-199-3

Query Match 1.9%; Score 74.2; DB 10; Length 2505;
Best Local Similarity 55.1%; Pred. No. 1.4e-09;
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 1806 ACACCACTTCATTTGCGAGCTGGGTATACAGAGTGTCGGTGGTGAATATCTGCTACAG 1865
DB 406 ACTGCCCTCCATATTGCTACAATAGCTGCCACCTAGAGGCTGCTGATGCTGTGGCAA 465
QY 1866 CATGGAGCTGATGTCATGCTTAAGATAAAGAGGCGCTTTGACCTTGCACAATGCATGT 1925
DB 466 CATGGAGCTGATGTCATGCTTAAGATAAAGAGGCGCTTTTTCACCTCCATTCGATTCGACGG 525
QY 1926 TCTTATGGACATTTATGAAGTTCGCAAACTTCTTGTAAACATGGAGCAGTAGTAAATGTA 1985
DB 526 TACTATGGACATGAACAGGTAACCTCGCCCTCTTTTGAATTTGGTCTGATGTAATGTA 585
QY 1986 GCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAAATATGAAATT 2045
DB 586 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATATT 645
QY 2046 TGCAAACTTCTGCTCCAGCATGG 2068
DB 646 GCAAACTCTTGATGGAGAGG 668

RESULT 15
US-09-947-199-1
Sequence 1, Application US/09947199
Patent No. US20020127684A1
GENERAL INFORMATION:
APPLICANT: Raju, Jeyaseelan
TITLE OF INVENTION: NOVEL CARX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: MNI-068CP2
CURRENT APPLICATION NUMBER: US/09/947,199
PRIOR FILING DATE: 2001-09-05
PRIOR APPLICATION NUMBER: 60/111,938
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/291,839
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: 09/458,457
PRIOR FILING DATE: 1999-12-10
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 3025
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS

LOCATION: (48)...(2552)
US-09-947-199-1

Query Match 1.9%; Score 74.2; DB 10; Length 3025;
Best Local Similarity 55.1%; Pred. No. 1.7e-09;
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 1806 ACACCACTTCATTTGCGAGCTGGGTATACAGAGTGTCGGTGGTGAATATCTGCTACAG 1865
DB 453 ACTGCCCTCCATATTGCTACAATAGCTGCCACCTAGAGGCTGCTGATGCTGTGGCAA 512
QY 1866 CATGGAGCTGATGTCATGCTTAAGATAAAGAGGCGCTTTGACCTTGCACAATGCATGT 1925
DB 513 CATGGAGCTAATGTCATTAATTAAGATGCGCTTTTTCACCTCCATTCGATTCGACGG 572
QY 1926 TCTTATGGACATTTATGAAGTTCGCAAACTTCTTGTAAACATGGAGCAGTAGTAAATGTA 1985
DB 573 TACTATGGACATGAACAGGTAACCTCGCCCTCTTTTGAATTTGGTCTGATGTAATGTA 632
QY 1986 GCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAAATATGAAATT 2045
DB 633 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATATT 692
QY 2046 TGCAAACTTCTGCTCCAGCATGG 2068
DB 693 GCAAACTCTTGATGGAGAGG 715

Search completed: February 12, 2003, 03:32:11
Job time : 215.419 secs


```

FEATURES               Helix Research Institute.
source                 Location/Qualifiers
1. .838
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE2000021"
/clone_lib="PLACE2"
/tissue_type="placenta"
/note="Vector: pME18SFL3"

BASE COUNT           260 a 143 c 203 g 227 t      5 others
ORIGIN

Query Match          20.6%; Score 784.2; DB 9; Length 838;
Best Local Similarity 98.0%; Pred. No. 5e-171;
Matches 823; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

QY 2734 CAGCCAGCAGTCTTGACAACTTATCTGGAGTCTTTCAGAACTGCTCTCAGTAGTATTAGTT 2793
|||||
Db 1 CAGCCAGCAGTCTTGACAACTTATCTGGAGTCTTTCAGAACTGCTCTCAGTAGTATTAGTT 60

QY 2794 CAAGTGGAAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTATA 2853
|||||
Db 61 CAAGTGGAAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTATA 120

QY 2854 GCATAACTCAATTCGTAAAGAACTCTTGACCTTGAGCACCCTAATGGATATATTTGAGACAG 2913
|||||
Db 121 GCATAACTCAATTCGTAAAGAACTCTTGACCTTGAGCACCCTAATGGATATATTTGAGACAG 180

QY 2914 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACACAGGAGCTCAAGGAGATTGGAA 2973
|||||
Db 181 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACACAGGAGCTCAAGGAGATTGGAA 240

QY 2974 TCAATGCTTATGGACATAGGCACAACTAATTAAGAGAGTCGAGAGACTTATCTCCGGAC 3033
|||||
Db 241 TCAATGCTTATGGACATAGGCACAACTAATTAAGAGAGTCGAGAGACTTATCTCCGGAC 300

QY 3034 AACAAAGTCTTAACCCATATTACTTTGAACACCTCTGCTAGTGGAACTTATCTTATAG 3093
|||||
Db 301 AACAAAGTCTTAACCCATATTACTTTGAACACCTCTGCTAGTGGAACTTATCTTATAG 360

QY 3094 ATCTGCTCTCTGATGATTAAGAGTTTTCAGTCTGTGGAGGAAGAGATGCAAGTACAGTTC 3153
|||||
Db 361 ATCTGCTCTCTGATGATTAAGAGTTTTCAGTCTGTGGAGGAAGAGATGCAAGTACAGTTC 420

QY 3154 GAGACACAGATGAGGTATGTCAGGTGGAATCTTCAACAGATACAAATTTCTCAAGA 3213
|||||
Db 421 GAGACACAGATGAGGTATGTCAGGTGGAATCTTCAACAGATACAAATTTCTCAAGA 480

QY 3214 TTCAGAAAGTTTGTACAGAAACTATGGAAAGATACACTCACCGGAGAAAGAGCTTT 3273
|||||
Db 481 TTCAGAAAGTTTGTACAGAAACTATGGAAAGATACACTCACCGGAGAAAGAGCTTT 540

QY 3274 CTGAAGAAACACACACATGCGCAATGAACGATGCTATTTCATGGTCTCTCTTTGTGA 3333
|||||
Db 541 CTGAAGAAACACACACATGCGCAATGAACGATGCTATTTCATGGTCTCTCTTTGTGA 600

QY 3334 ATGCAATTTATCCCAAGGCTTTGATGAAGGATCGGTACATAGTGGTATGTTTGGAG 3393
|||||
Db 601 ATGCAATTTATCCCAAGGCTTTGATGAAGGATCGGTACATAGTGGTATGTTTGGAG 660

QY 3394 CTGGCAATTTATTTGCTGAAAACCTCTTCCAAAAGCAATCAATATGATATGGAATTTGGAG 3453
|||||
Db 661 CTGGCAATTTATTTGCTGAAAACCTCTT - CAAGAAGCAATCAATATGATATGGAATTTGGAN 719

QY 3454 GAGGTACTGGGTGTCAGTTTCAAAAGACAGATCTTGTGTACATTTGCCACAGCAGCTGC 3513
|||||
Db 720 GAGT -CTGGGTGTCAGTTTCAAAAGACAGCTTGTGTACATTTGCCACAGCAGCTGC 778

QY 3514 TCTTTTGTG -CCGGTAACTCTGGAAAGTCTTCTCTGCAAGTTCAGTGCATCAAAATGCG 3571
|||||
Db 779 TCTTTTGGCCGGGTAAACCTCTGGAAAGTCTTCTCTGCAAGTTCANTGCAATGAAATGCG 838
|||||

```

RESULT 2
 BM457025
 LOCUS
 DEFINITION
 1128 bp mRNA linear EST 05-FEB-2002
 5', mRNA sequence.
 BM457025
 BM457025.1 GI:18506065
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1128)
 NIH-MGC http://mgs.nci.nih.gov/.
 JOURNAL
 COMMENT
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12346 Row: a column: 18
 High quality sequence stop: 507.
 Location/Qualifiers
 1. .1128
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5583257"
 /clone_lib="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /notes="Organ: testis; Vector: pCMV-SPORT6; Site: 1; NotI;
 Site: 2; SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH-MGC Library."
 BASE COUNT 328 a 297 c 232 g 271 t
 ORIGIN
 Query Match 17.6%; Score 672.2; DB 13; Length 1128;
 Best Local Similarity 97.7%; Pred. No. 6.1e-145;
 Matches 724; Conservative 0; Mismatches 13; Indels 4; Gaps 4;
 QY 1976 AGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAA 2035
 |||||
 Db 1 AGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAA 60
 QY 2036 ATATGAAATTTGCAAACTCTGCTCCAGCATGGTCAGACCCCTACCAAAAAAACAGGGA 2095
 |||||
 Db 61 ATATGAAATTTGCAAACTCTGCTCCAGCATGGTCAGACCCCTACCAAAAAAACAGGGA 120
 QY 2096 TGAATACTCTCTTTGGATCTTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAG 2155
 |||||
 Db 121 TGAATACTCTCTTTGGATCTTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAG 180
 QY 2156 GGCAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTAGCCAGAGTGAAGAAGTT 2215
 |||||
 Db 181 GGCAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTAGCCAGAGTGAAGAAGTT 240
 QY 2216 GTCTTCTCTGATATGTAATGTAATGCGCGATACCCAGGACAGACATTTCAACACCTTTTACA 2275
 |||||
 Db 241 GTCTTCTCTGATATGTAATGTAATGCGCGATACCCAGGACAGACATTTCAACACCTTTTACA 300
 QY 2276 TTTAGCAGCTGGTATAATATTTAGAAAGTTCAGAGTATTTGTTTACAACACGAGCTGA 2335
 |||||
 Db 301 TTTAGCAGCTGGTATAATATTTAGAAAGTTCAGAGTATTTGTTTACAACACGAGCTGA 360
 QY 2336 TGTGAATGCCAAGACAAAGGAGGACTTATTCCTTTACATATATGACAGATCTTACGGGCA 2395
 |||||


```
source
1. .691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4537138"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DHL0B (phase-resistant)"
/site_1="Organ: testis; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 218 a 122 c 165 g 186 t
ORIGIN
Query Match 16.9%; Score 643.8; DB 12; Length 691;
Best Local Similarity 99.3%; Pred. No. 1.8e-138;
Matches 678; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY 1725 AAGCGTGAGATGTCGAACCTGTAATAAAGTGTCTACTGTCAGAGTGTCAACTGCAGA 1784
Db 10 AAGCGTGAGATGTCGAACCTGTAATAAAGTGTCTACTGTCAGAGTGTCAACTGCAGA 69
QY 1785 GACATTTGAAGGCGTCAGCTACACCACTTCATTTTGCAGCTGGGTATACAGAGTGTC 1844
Db 70 GACATTTGAAGGCGTCAGCTACACCACTTCATTTTGCAGCTGGGTATACAGAGTGTC 129
QY 1845 GTGTGGAATATCTCTACAGCATGGAGCTGATGTGATGCTAAAGATAAAGAGGCGCTT 1904
Db 130 GTGTGGAATATCTCTACAGCATGGAGCTGATGTGATGCTAAAGATAAAGAGGCGCTT 189
QY 1905 GTACCTTTGCACATGCATGTTCTTATGGACATTTATGAAGTTGCAGACCTTCTTTTAA 1964
Db 190 GTACCTTTGCACATGCATGTTCTTATGGACATTTATGAAGTTGCAGACCTTCTTTTAA 249
QY 1965 CATGAGCAGTAGTAAATGATGAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCA 2024
Db 250 CATGAGCAGTAGTAAATGATGAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCA 309
QY 2025 GCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCA-TGGTGCAGACCCCTACCA 2083
Db 310 GCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCACTGGTGCAGACCCCTACCA 369
QY 2084 AAAAAACAGGATGGAATATCTCTTTGGATCTTTTAAAGATGGAGATACAGATATTC 2143
Db 370 AAAAAACAGGATGGAATATCTCTTTGGATCTTTTAAAGATGGAGATACAGATATTC 429
QY 2144 AGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAG 2203
Db 430 AGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAG 489
QY 2204 AGTGAAGAAGTTGCTTCTCTGATA-ATGTAATTTGCCGCGATACCCAAAGGCAGACATT 2262
Db 490 AGTGAAGAAGTTGCTTCTCTGATAATGTAATTTGCCGCGATACCCAAAGGCAGACATT 549
QY 2263 CAACACCTTTACATTTAGCAGCTGGTATATTAATTTAGAGTTGCGAGTATTTGTTAC 2322
Db 550 CAACACCTTTACATTTAGCAGCTGGTATATTAATTTAGAGTTGCGAGTATTTGTTAC 608
QY 2323 AACACGAGCTGATGTGAATGCCCAAGACAAAGAGGACCTTATTCCTTTACATAATGCAG 2382
Db 609 AACACGAGCTGATGTGAATGCCCAAGACAAAGAGGACCTTATTCCTTTACATAATGCAG 668
QY 2383 CATCTTACGGGCATGTAGATGTA 2405
Db 669 CATCTTACGGGCATGTAGATGTA 691
RESULT 5
AL601027
LOCUS
DEFINITION
DKFp313G2239_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
759 bp mRNA linear EST 14-AUG-2001
DKFp313G2239 5', mRNA sequence.
```

```
AL601027
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 759)
Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
,S.
EST (Duesterhoeft, et al.)
Unpublished (1999)
Contact: Duesterhoeft A
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Oligen (Hilden/Germany) within the cDNA sequencing
consortium of the German genome Project.
No sl sequence available.
This clone (DKFp313G2239) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
source
1..759
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFp313G2239"
/clone_lib="313 (synonym: hlcc2)"
/dev_stage="adult"
/lab_host="DHL0B"
/note="vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"
```

```
BASE COUNT 217 a 156 c 175 g 201 t 10 others
ORIGIN
Query Match 16.8%; Score 642.8; DB 9; Length 759;
Best Local Similarity 97.2%; Pred. No. 3.3e-138;
Matches 669; Conservative 0; Mismatches 17; Indels 2; Gaps 2;
QY 2393 GCATGTAGATGTAGCAGCTTACTATAAAGTAAATGCTGTCATGCGACGGACAA 2452
Db 72 GCATGTANATGTAGCAGCTTACTATAAAGTAAATGCTGTCATGCGACGGACAA 131
QY 2453 ATGGCGTTTACACCTTTGCAGAGCAGGCCCAAAAGGAGACACAGCTTTTGTGCTT 2512
Db 132 ATGGCGTTTACACCTTTGCAGAGCAGGCCCAAAAGGAGACACAGCTTTTGTGCTT 191
QY 2513 GTTGTAGCCCATGGAGCTGACCGGACTCTTAAATAATCAGGAAGGACACACCTTTTGA 2572
Db 192 NTTTTGGCCCATGGAGCTGACCGGACTCTTAAATAATCAGGAAGGACACACCTTTTGA 251
QY 2573 TTTAGTTTCAGCGGATGATGTACGCGCTTCTTGACAGCAGCCATGCCCATCTGCTCT 2632
Db 252 TTTAGTTTCAGCGGATGATGTACGCGCTTCTTGACAGCAGCCATGCCCATCTGCTCT 311
QY 2633 GCCTCTTGTTCAGCGCTCAAGTGTCTCAATGTTGTGAGAGCCCGAGGAGCCATGCGA 2692
Db 312 GCCTCTTGTTCAGCGCTCAAGTGTCTCAATGTTGTGAGAGCCCGAGGAGCCATGCGA 371
QY 2693 TGCTCTCTTTCAGGTCCTCATCAGCCCATCAGCCCTTCTGACGCGCAGCTTTTGACAA 2752
Db 372 TGCTCTCTTTCAGGTCCTCATCAGCCCATCAGCCCTTCTGACGCGCAGCTTTTGACAA 431
QY 2753 CTTATCTGGAGTTTTCAGAACTGCTTTCAGTAGTTAGTTCAAGTGAACAGAGGTCG 2812
Db 432 CTTATCTGGAGTTTTCAGAACTGCTTTCAGTAGTTAGTTCAAGTGAACAGAGGTCG 491
QY 2813 TTCACCTTTGGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATACCTCAATTCGTAAG 2872
Db 492 TTCACCTTTGGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATACCTCAATTCGTAAG 551
```

```

QY 2873 GAATCTTGAGCTTGGACACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGT 2932
|||||
Db 552 GAATCTTGAGCTTGGACACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGT 611
|||||
QY 2933 ATTAGTTGAGATGGGACAGAGAGCTGAAGAGGATTTGGAATCAATGCTTTATGGACATAG 2992
|||||
Db 612 ANNAGTTGAGATGGGACAGAGAGCTGAAGAGGATTTGGAATCAATGCTTTATGGACATAG 671
|||||
QY 2993 GCACAACTAATTTAAA-GGAGTCGAGAGACTTATCTCCGCGACACAA-GGTCTTAACCCA 3050
|||||
Db 672 GCACAACTAATTTAAAAGGAGTCGAGAGACTTATCTCCGCGACACAAAGGGGCTTAACCCA 731
|||||
QY 3051 TATTAACTTTGAACACTCTGGTAGTG 3078
|||||
Db 732 TATTAACTTTGAACACTCTGGNAGAG 759
|||||

RESULT 6
BO885764
LOCUS
DEFINITION AGENCOURT_8746924 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6392888
5', mRNA sequence.
ACCESSION BO885764
VERSION BO885764.1 GI:22277782
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 950)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13883 row: h column: 09
High quality sequence stop: 641.
FEATURES
source
1..950
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6392888"
/clone.lib="NIH_MGC_130"
/lab_host="DH10B (phage-resistant)"
/note="Organ: otocysts; Vector: pCMV-SPORT6.1.cdb;
Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally.
Primer: Oligo dT. Average insert size 1.95 kb.
Constructed by ResGen, Invitrogen Corp. Note: this is a
NIH_MGC Library."
BASE COUNT 280 a 195 c 226 g 245 t 4 others
ORIGIN
Query Match 16.6%; Score 634.4; DB 14; Length 950;
Best Local Similarity 86.4%; Pred. No. 3.4e-136;
Matches 759; Conservative 0; Mismatches 111; Indels 8; Gaps 5;

QY 2500 AGCTTTGTGCTTGTGCTAGCCCATGGAGCTGACCCGACTCTTAAATCATCAGGAGGAC 2559
|||||
Db 1 AGCTTTGTGCTTGTGCTAGCCCATGGAGCTGACCTACTCTT-AAATCATCAGGAGGAC 59
|||||
QY 2560 AAACACCTTTAGATTTTACGGGATGATGTCAGCCGCTCTTCTGACAGCAGCCATGC 2619
|||||
Db 60 AAACACCTTTAGATTTTACGGGATGATGTCAGCTGCTGACAGCAGCCATGC 119
|||||
QY 2620 CCCCATCTGCTGCTGCTTGTGTACAGGCTCAAGTGTCTCAATGGTGTGAGAGCCGAC 2679
|||||

```

```

Db 120 CCCCTCTCTGCTGCTACGTGCTACAAACCTCAAGTGTGCTGAGGGGCCCGG 179
|||||
QY 2680 GAGCCACTCCAGATGCTCTCTCTCAGTCCATCTAGCCCATCAAGCCCTTTCTGCGAGCA 2739
|||||
Db 180 GAGCCACTCCAGATGCTCTCTCAGTCCATCTAGCCCATCAAGCCCTTTCTGCGAGCA 239
|||||
QY 2740 GCAGTCTTGACAACTTATCTGGGAGTTTTCAGAACTGTCTTCAGTAGTGTCAAGTG 2799
|||||
Db 240 GCAGCTTCACAACTTATCTGGCAGCTTCTCGAACTGTCCGAGTGGTGTAGTTCAGTG 299
|||||
QY 2800 GAACAGAGGCTGCTTCCAGTTTGGAGAAAAGAGGTTCCAGAGGTAGATTTTACATAA 2859
|||||
Db 300 CAGCAGAAGGCTGCTACTGTTTGCAAAAGAAAGAGAGATTTCAGGAATTCGATTTAGTATAA 359
|||||
QY 2860 CTCATTCGTGAAGAACTTGGACTTGCACCTAATGATATATTTGAGAGAGACAGA 2919
|||||
Db 360 CTCAGTTTCATAAGAACTTGGACTTGCAGCTTGCAGTATATTTGAGAGAGACAGA 419
|||||
QY 2920 TCACCTTTGGATGATTTAGTTGAGATGGGCGACAAAGGAGTGAAGGATTTGAATCATG 2979
|||||
Db 420 TTACCTTAGATGCTCTAGTTGAAATGGCCACAAAGAACTTGAAGAGATTGGAATCAATG 479
|||||
QY 2980 CTTATGCATAGGACAACTAATTTAAAGGACTCGAGAGACTTATCTCGGACACAAAG 3039
|||||
Db 480 CCTATGSACATCGACACAAAGCTGATTAAAGGAGTTGAAAGGCTGATCTCTGGACACAAAG 539
|||||
QY 3040 GTCTTAACCCATATTTAACTTTGAACACCTCTGTTAGTGAACAAATTTCTTATAGATCTGT 3099
|||||
Db 540 GTCTTAACCCATATTTAACTTTGAACACCTCTGTTAGTGAACAAATTTCTTATAGATCTGT 599
|||||
QY 3100 CTCTGATGATAAAGAACTTTCAGTCTGTGGAGGAGAGATGCAAAAGTACAGTTGCGAGAGC 3159
|||||
Db 600 CTCTGATGATAAAGAACTTTCAGTCTGTGGAGGAGAGATGCAAAAGTACAGTTGCGAGAGC 659
|||||
QY 3160 ACAGAGATGAGGATCATGTCAGCTGGAATCTTCAACAGATACATATTTCTCAAGATTCAGA 3219
|||||
Db 660 ACAGAGATGAGGATCATGTCAGCTGGAATCTTCAACAGATACATATTTCTCAAGATTCAGA 719
|||||
QY 3220 AGTTTCTTAAACAACTATGGAAAGATACATCACTACCGG-AGAAAGAAAGTTTCTGAA 3278
|||||
Db 720 AGTTTCTTAAACAACTATGGAAAGATATACACACCGGAGAAAGAAAGTTTCTGAA 779
|||||
QY 3279 GAAACCCACACCATGC-CAATGAACGAATGCTATTTTCATGGGTCT---CCTTTTGTGAA 3334
|||||
Db 780 GAAACCCACACCATGC---CCTTTGATGAAAGGATCTTATTTTCATGGGTCTCCTTTTGTGAA 839
|||||
QY 3335 TGCATTTATCCACAAAGG--CTTTGATGAAAGGATGC 3370
|||||
Db 840 TGCATTTATCCAAAGGCTTTTGTGAAAGGATGC 877
|||||

RESULT 7
BO875327
LOCUS
DEFINITION RCI-BF0623-120200-011-d04 BT0623 Homo sapiens cDNA, mRNA sequence.
ACCESSION BO875327
VERSION BO875327.1 GI:14252306
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 648)
REFERENCE
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

```

MEDLINE
COMMENT

20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml.pl?tl=et2-RC1-BT0623-120
200-011-d004&t3=2000-02-12&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 4
High quality sequence stop: 648.

FEATURES
source

1..648
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT0623"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
/16 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 200 a 118 c 157 g 172 t 1 others

Query Match 15.9%; Score 605; DB 12; Length 648;

Best Local Similarity 98.7%; Pred. No. 1.8e-129;

Matches 620; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 1663 TCCTCCAAGAGGTATCTCATTAGTAATTCAGAGCAGACACAAATTCCTGGAGCTG 1722

Db 23 TCTCATAGAGGGTATCTCATTAGG--ATTCAGAGGCGAGACAAATTCCTGGAGCTG 80

QY 1723 CAAAGCTGGAGATGTCGAACTGTAAAAAACTGTACTGTTCAGAGTGTCAACTGCA 1782

Db 81 CAAAGCTGGAGATGTCGAACTGTAAAAAACTGTACTGTTCAGAGTGTCAACTGCA 140

QY 1783 GAGACATTGAAGGGCGCTCAGCTACACCACCTTCATTTCGAGCTGGGTATACAGAGTGT 1842

Db 141 GAGACATTGAAGGGCGCTCAGCTACACCCTTCATTTCGAGCTGGGTATACAGAGTGT 200

QY 1843 CCGTGTGGAATATCTGCTACAGCATGGAGCTGATGTCATGCTAAAGATAAAGGAGGCC 1902

Db 201 CCGTGTGGAATATCTGCTACAGCATGGAGCTGATGTCATGCTAAAGATAAAGGAGGCC 260

QY 1903 TTGTACCTTTGCACATGTCATGCTTTATGGACATTTATGAAGTTGCAGAACTCTTGTGTA 1962

Db 261 TTGTACCTTTGCACATGTCATGCTTTATGGACATTTATGAAGTTGCAGAACTCTTGTGTA 320

QY 1963 AACATGGACAGTATGTAATGATGATGATTTATGAAATTTACACCTTTACATGAAGCAG 2022

Db 321 AACATGGACAGTATGTAATGATGATGATTTATGAAATTTACACCTTTACATGAAGCAG 380

QY 2023 CAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTCAGACCCCTACCA 2082

Db 381 CAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTCAGACCCCTACCA 440

QY 2083 AAAAAACAGGATGGAATATCTCTTTGGATCTTTGATGTTAAAGATGGAGATACAGATATTC 2142

Db 441 AAAAAACAGGATGGAATATCTCTTTGGATCTTTGATGTTAAAGATGGAGATACAGATATTC 500

QY 2143 AAGATCTGCTTACGGGAGATGACGCTTTGCTAGATGCTGCCAAGAGGCTTTTAGCCA 2202

Db 501 AAGATCTGCTTACGGGAGATGACGCTTTGCTAGATGCTGCCAAGAGGCTTTTAGCCA 560

QY 2203 GAGTGAAGAAGTTGCTTCTCCTGATAATGTAAATTCGCCGCGATACCCAAAGCGACATTT 2262

Db

561 GAGTGAAGAAGTTGCTTCTCCTGATAATGTAATTCGCGGATACCCAGGACACATT 620

QY

2263 CAACACCTTTACATTTAGCAGCTGGTTA 2290

Db

621 CAACACCTTTACATATAGCAGCTGGTTA 648

RESULT 8

BE220005

BE220005/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 622)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-femail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov

Seq primer: -40UP from Gibco

High quality sequence stop: 474.

Location/Qualifiers

1..622

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NCI-CGAP_Lu24"

/tissue_type="carcinoid"

/lab_host="DH10B"

/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a

modified polylinker; plasmid DNA from the normalized

library NCI-CGAP_Lu5 was prepared, and ss circles were

made in vitro. Following HAP purification, this DNA was

used as tracer in a subtractive hybridization reaction.

The driver was PCR-amplified cDNAs from a pool of 5,000

clones made from the same library (cloneIDs

1414920-1417991 and 1520904-1522439). Subtraction by Bento

Soares and M. Fatima Bonaldo."

BASE COUNT 179 a 130 c 128 g 185 t

ORIGIN

Query Match 15.5%; Score 593.2; DB 10; Length 622;

Best Local Similarity 97.1%; Pred. No. 9.8e-12;

Matches 604; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 3183 GGAATCTTCAACAGATACATAATTTCTCAAGATTCAGAGGTTTGTAAACAGAACTATCG 3242

Db 622 GCAATCTTCAACAGTACTATATTTCTCAAGATTCAGAGGTTTGTAAACAGAACTATCG 563

QY 3243 GAAATACACTCCGCGGAGAAAGAGTTTCTGAAGAAACCAACCATGCCAATGAA 3302

Db 562 CAAATACACTCCGCGGAGAAAGAGTTTCTGAAGAAACCAACCATGCCAATGAA 503

QY 3303 CCAATGCTATTTCATGGGTCTCTCTTTTGTGAATCAATATCCACAAAGGCTTTGTATGA 3362

Db 502 CGAATGCTATTTCATGGGTCTCTCTTTTGTGAATCAATATCCACAAAGGCTTTGTATGA 443

QY 3363 AGGCATCGGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTCTGGAACACTCTTCC 3422
 Db 442 AGGCATCGGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTCTGGAACACTCTTCC 383
 QY 3423 AAAAGCAATCAATATGATATGGAATTTGGAGAGGTACTGGGTGTCAGTTCACAAAGAC 3482
 Db 382 AAAAGCAATCAATATGATATGGAATTTGGAGAGGTACTGGGTGTCAGTTCACAAAGAC 323
 QY 3483 AGATCTGTTTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTCT 3542
 Db 322 AGATCTGTTTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTCT 263
 QY 3543 TTCTGTCAGTTCAGTGAATGGAATGCAATCTCTCCAGGTGATCACTCAGTCACT 3602
 Db 262 TTCTGTCAGTTCAGTGAATGGAATGCAATCTCTCCAGGTGATCACTCAGTCACT 203
 QY 3603 GGTAGGCCAGTGAATGGCTAGCTATAGCTGAATATGTTATTTACAGAGGAGAACAG 3662
 Db 202 GGTAGGCCAGTGAATGGCTAGCTATAGCTGAATATGTTATTTACAGAGGAGAACAG 143
 QY 3663 GCTATCTCAGTATTTAATTTACTTACAGATTTAGGCGCTGAGGTATGGTTCGATGGA 3722
 Db 142 GCTATCTCAGTATTTAATTTACTTACAGATTTAGGCGCTGAGGTATGGTTCGATGGA 83
 QY 3723 TAAATAGTTATTTTAAAGAACTAATTCACCTGAACCTTAAATCATCAAAAGCAGAGTGGC 3782
 Db 82 TAAATAGTTATTTTAAAGAACTAATTCACCTGAACCTTAAATCATCAAAAGCAGAGTGGC 23
 QY 3783 CTCTACGTTTACTCCTTTGCT 3804
 Db 22 CTCTACGTTTACTCCTTTGCT 1

RESULT 9
 BE467629/c
 LOCUS
 DEFINITION
 ADP-RIBOSE POLYMERASE. ; mRNA sequence.
 BE467629
 BE467629.1 GI:9513404
 EST.
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 616)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapsb-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
 Seq primer: -400P from Gibco
 High quality sequence stop: 402.
 Location/Qualifiers
 1..616
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3213021"
 /clone_lib="NCI-CGAP_Lu24"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a

modified polylinker; Plasmid DNA from the normalized
 library NCI-CGAP_Lu5 was prepared, and ss circles were
 made in vitro. Following HAP purification, this DNA was
 used as tracer in a subtractive hybridization reaction.
 The driver was PCR-amplified cDNAs from a pool of 5,000
 clones made from the same library (clonoids
 1414920-1417991 and 1520904-1522439). Subtraction by Bento
 Soares and M. Fatima Bonaldo. "

BASE COUNT 179 a 130 c 124 g 183 t
 ORIGIN

Query Match 15.3%; Score 584; DB 10; Length 616;
 Best Local Similarity 96.8%; Pred. No. 1.3e-124;
 Matches 596; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 3189 TTCACAGATACATATTTCTCAAGATTTCAGAGTTTGTAAACAAGAACTATGGGAAGA 3248
 Db 616 TTCACAGATACATATTTCTCAAGATCATATAGTTTGCACCACTATGGATAGA 557
 QY 3249 TACACTCACCGGAGAAAAGATTTCTGAAGAAAACCAACCATGCCAATGAACGAATG 3308
 Db 556 TACACTCACCGGAGACAAGAAGTTTGTGAAGAAAACCCACCATGCCAATGACGAATG 497
 QY 3309 CTATTTTCATGGTCTCCTTTTGTGATGCAATTTATCCAAAAGGCTTTGATGAAGGCAT 3368
 Db 496 CTATTTTCATGGTCTCCTTTTGTGATGCAATTTATCCAAAAGGCTTTGATGAAGGCAT 437
 QY 3369 GCCTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTCTGAAACCTTTCCAAAAGC 3428
 Db 436 GCCTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTCTGAAACCTTTCCAAAAGC 377
 QY 3429 AATCAATATGATATGGAATTTGGAGAGGACTGGGTGTCAGTTTCCAAAAGCAGATCT 3488
 Db 376 AATCAATATGATATGGAATTTGGAGAGGACTGGGTGTCAGTTTCCAAAAGCAGATCT 317
 QY 3489 TGTTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTTGGAAAGTCTTTCTCG 3548
 Db 316 TGTTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTTGGAAAGTCTTTCTCG 257
 QY 3549 CAGTTTCACTGCAATGAAATGGCAGATTTCTCTCCAGGTGATCATCAGTCACTGCTAGG 3608
 Db 256 CAGTTTCACTGCAATGAAATGGCAGATTTCTCTCCAGGTGATCATCAGTCACTGCTAGG 197
 QY 3609 CCCAGTGAATGCCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGACAGGCTTAT 3668
 Db 196 CCCAGTGAATGCCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGACAGGCTTAT 137
 QY 3669 CCTGAGTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3728
 Db 136 CCTGAGTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 77
 QY 3729 GTTATTTTAAAGAACTAATTTCCACTGAACCTTAAATTAATTAATTAATTAATTAATTAAT 3788
 Db 76 GTTATTTTAAAGAACTAATTTCCACTGAACCTTAAATTAATTAATTAATTAATTAATTAAT 17
 QY 3789 GTTTTACTCCTTTGCT 3804
 Db 16 GTTTTACTCCTTTGCT 1

RESULT 10
 BE910901
 LOCUS
 DEFINITION
 601661978F1 NCI-CGAP_Mam1 Mus musculus cDNA clone IMAGE:3962251 5',
 mRNA sequence.
 BE910901
 BE910901.1 GI:10407885
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus
 house mouse.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 956)
 REFERENCE

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-femail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM9128 row: k column: 20
 High quality sequence stop: 672.

FEATURES
 source
 1..956
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:3962251"
 /clone_lib="NCI CGAP Maml"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"
 BASE COUNT 292 a 183 c 238 g 242 t 1 others
 ORIGIN

Query Match 14.9%; Score 570.4; DB 12; Length 956;
 Best Local Similarity 85.2%; Pred. No. 2.5e-121;
 Matches 673; Conservative 0; Mismatches 111; Indels 6; Gaps 3;

QY 2847 GATTTAGCATACCAATTCGTAAGCAATCTGGACTTGGACCACTAATGGATATATTT 2906
 |||||
 Db 1 GATTTAGTATACCTCAGTTCTAAGGAACCTTGGACTTGGACCACTAATGGATATATTT 60

QY 2907 GAGAGAGAACACATCACTTTGGATGTTAGTTGACATGGGGCACAAAGGAGCTGAAGGAG 2966
 |||||
 Db 61 GAGAGAGAACACATCACTTTAGTATGCTTGTGTAATGGCCCAAGCACTGAAGAG 120

QY 2967 ATTGGGAATCAATGCTTATGAGATAGGACCAACCAATTAAGGAGTCGAGAGACTTATC 3026
 |||||
 Db 121 ATTGGGAATCAATGCTTATGAGATAGGACCAACCAATTAAGGAGTCGAGAGACTTATC 180

QY 3027 TCCGGACACAAAGGCTTTAAACCCATATTTAACTTTGAACACCTCTGTTAGTGGACAAT 3086
 |||||
 Db 181 TCTGGACACAAAGGCTTTAACTTTGAACACCTCTGTTAGTGGACAAT 240

QY 3087 CTTATAGATCTCTCTCGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAAGT 3146
 |||||
 Db 241 CTTATAGATCTCTCTCGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAGT 300

QY 3147 ACAGTTCGAGACACAGAGATGAGTCTGATGAGTGGATCTTCAACAGATACAAATATT 3206
 |||||
 Db 301 ACTGTGCGGGACACAGAGATGGTGGTCACGCGCGCGCTTCAACAGATACAAATATT 360

QY 3207 CTCAGATTCAGAAGTTTGTAAAGAAACCTATGCGGAAAGATACACTCACCGGAGAAA 3266
 |||||
 Db 361 CTCAGATTCAGAAGTTTGTAAAGAAACCTATGCGGAAAGATACACTCACCGGAGAAA 420

QY 3267 GAAGTTTCTGAAGAAACCAACCAATGCGCAATGCAATGCAATGCTATTTTCATGGGTCTCT 3326
 |||||
 Db 421 GAAGTTTCTGAAGAAACCAACCAATGCGCAATGCAATGCTATTTTCATGGGTCTCT 480

QY 3327 TTTGTGAATGCAATTTATCCAAAGGCTTTGATGAAGGCAATGCTATGATGCTGATG 3386
 |||||
 Db 481 TTTGTGAATGCAATTTATCCAAAGGCTTTGATGAAGGCAATGCTATGATGCTGATG 540

QY 3387 TTTGAGCTGGCAATTTATTTTCTGAAACCTCTTCCAAAGCAATCAATATGATATGCA 3446
 |||||

Db 541 TTTGGAGCTGGAATTTATTTTGTGAGAACTCTTCCAAAAGCAATCAATATGTGTATGGA 600
 QY 3447 ATTGGAGGAGG--TACTGGGTGTCAGTTTACAAAACACAGATCTTGTACATTTGCCAC 3503
 |||||
 Db 601 ACTTGGAGTGGCACCAGGATGTCCAATTCACAAAACACAGATGCTGTACCTTTGTACAC 660
 |||||
 QY 3504 AGGAGCTGCTCTTTTCCGGGTAACTTGGGAAAGTCTTTTCCTGCAATGCAATG 3563
 |||||
 Db 661 AGGAGCTGCTGTTTTCGTCGAGT-ACCTTGGGCAAGTCTTTTCTTGGCCCTTCAACGCAATG 719
 |||||
 QY 3564 AAAATGACCAATCTCTCCAGGTCATCACTCAGTCACTGG--TAGGCCCACTGTAATG 3621
 |||||
 Db 720 ACATGACCAATCTCTCCAGGTCATCACTCAGTCACTGG--TAGGCCCACTGTAATG 779
 |||||
 QY 3622 GCCTAGCAAT 3631
 |||||
 Db 780 GCCTAGCAAT 789

RESULT 11
 BB339554
 LOCUS
 DEFINITION
 BB339554
 ACCESSION
 BB339554
 VERSION
 BB339554.2
 GI:16404413
 KEYWORDS
 EST.
 SOURCE
 house mouse.
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 702)
 Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
 Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
 Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,F., Toyota,T.,
 Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 Unpublished (2001)
 On Jul 11, 2000 this sequence version replaced gi:9048317.
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,
 M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,F., Matsura,
 S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
 Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,
 Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,
 K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
 Hayashizaki,Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, V., Williams, I., Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)

TITLE JOURNAL COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

Tel: 617-495-1812
Fax: 617-495-8557

Email: dmelton@bioh.harvard.edu

Library was constructed by Dr. J. J. Ferrer in vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Seq primer: -40Rp from Gibco
High quality sequence stop: 429.

FEATURES source

Location/Qualifiers
1. .553
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE: 5778810"
/clone_lib="Human insulinoma"
/tissue_type="insulinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pBluescript SK-; Site_1: xhoI; Site_2: EcoRI; Constructed with lambda ZAPII system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol
(http://genome.wustl.edu/est/Lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."

BASE COUNT
ORIGIN

Query Match 14.4%; Score 549; DB 14; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.7e-116;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3260 GAGAAAGAGAGTTCTTGAAGAAACACACACCATGCCAATGAACGAATGCTATTTCATGG 3319
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5 GAGAAAGAGAGTTCTTGAAGAAACACACACCATGCCAATGAACGAATGCTATTTCATGG 64
QY 3320 GTCTCCTTTTGTGAATGCATTTATCCACAAAGGCTTTGTATGAAGGCATCGGTACATAGG 3379
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
65 GTCTCCTTTTGTGAATGCATTTATCCACAAAGGCTTTGTATGAAGGCATCGGTACATAGG 124
QY 3380 TGGTATGTTTGGAGCTGGCATTATTTTGTGTAAGAACTCTTCCAAAGCAATCAATATGT 3439
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
125 TGGTATGTTTGGAGCTGGCATTATTTTGTGTAAGAACTCTTCCAAAGCAATCAATATGT 184
QY 3440 ATATGGAATTTGGAGAGGAGTACTGGGTGTCAGTTTCACAAAGACAGATCTTTGTACATTG 3499
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
185 ATATGGAATTTGGAGAGGAGTACTGGGTGTCAGTTTCACAAAGACAGATCTTTGTACATTG 244
QY 3500 CCACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCTTTCCTGCGAGTTTCAGTGC 3559
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
245 CCACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCTTTCCTGCGAGTTTCAGTGC 304
QY 3560 AATGAAATGGCAGATCTTCTCCAGGTCTCATCTCAGTCAGTGGTGGAGGAGTGTAA 3619
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
305 AATGAAATGGCAGATCTTCTCCAGGTCTCATCTCAGTCAGTGGTGGAGGAGTGTAA 364
QY 3620 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 3679
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
365 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 424

QY 3680 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCCGATGATAAATAGTTATTTTAAG 3739
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
425 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCCGATGATAAATAGTTATTTTAAG 484
QY 3740 AAACCTAATTCCCACTGAACCTAAATCATCAAAAGCAGCAGTGGCCTCTAGGTTTACTCTCT 3799
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
485 AAACCTAATTCCCACTGAACCTAAATCATCAAAAGCAGCAGTGGCCTCTAGGTTTACTCTCT 544
QY 3800 TTGCTGAAA 3808
Db ||||||||
545 TTGCTGAAA 553

Search completed: February 11, 2003, 21:15:50
Job time : 4991.35 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:31:44 ; Search time 37.6068 Seconds
(without alignments)
3897.581 Million cell updates/sec

Title: US-09-843-159B-3
Perfect score: 5769
Sequence: 1 GFGKDDVVEYLLONGASVQA.....AYPEYLITYQIMRPGCMVDG 1100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	5769	100.0	1100	22	Tankyrase homology
2	5766	99.9	1166	22	Human tankyrase2 T
3	5766	99.9	1169	22	Human tankyrase2 I
4	5766	99.9	1169	22	Human tankyrase2 C
5	5766	99.9	1240	22	Tankyrase homology
6	5766	99.9	1262	22	Human tankyrase2 C
7	5766	99.9	1385	22	Human tankyrase2 T
8	5750	99.7	1166	21	Human tankyrase II
9	5717	99.1	1166	22	Human tankyrase ho
10	5605	97.2	1074	20	Human Grb7 effecto

11	5201	90.2	1166	22	Human SPANK. Homo
12	4932.5	85.5	1333	21	Human tankyrase II
13	4889.5	84.8	1327	21	Human tankyrase I
14	4889.5	84.8	1327	21	Human tankyrase.
15	4889.5	84.8	1327	22	Human tankyrase1 S
16	4889.5	84.8	1327	23	Human tankyrase.
17	4859	84.2	1431	21	Human tankyrase II
18	4703.5	81.5	991	22	Mouse SPANK. Mus
19	4624.5	80.2	907	22	Human breast cance
20	4102	71.1	784	22	Human tankyrase2 C
21	3959	68.6	756	22	Human tankyrase2 c
22	3835.5	66.5	1181	22	Drosophila melanog
23	3835.5	66.5	1181	22	Drosophila tankyra
24	3596	62.3	802	23	Protein of App rel
25	3093	53.6	949	21	Human truncated ta
26	3093	53.6	949	21	Human truncated ta
27	2367	41.0	522	22	Human tankyrase2 c
28	2165	37.5	415	22	Human Grb7 effecto
29	2053.5	35.6	1099	22	Human tankyrase2 e
30	1942	33.7	385	22	Human T cell leuka
31	1918.5	33.3	673	21	Human truncated ta
32	1918.5	33.3	673	23	Human TRFL TANK2 b
33	1895	32.8	362	22	Human breast cance
34	1706	29.6	331	22	Human protein sequ
35	1395	24.2	261	22	Human ORFX ORF3133
36	1251	21.7	240	21	Novel human diagno
37	875.5	15.2	4274	22	Novel human diagno
38	875.5	15.2	4386	22	Human tankyrase II
39	874	15.1	160	21	Human tankyrase I
40	874	15.1	160	22	Human tankyrase I
41	860.5	14.9	1872	22	Human secreted pro
42	860	14.9	193	22	Human tankyrase I
43	852.5	14.8	4397	22	Novel human diagno
44	836	14.5	1549	22	Drosophila melanog
45	836	14.5	1549	22	Drosophila melanog

ALIGNMENTS

RESULT 1
AA97748
ID AA97748 standard; Protein; 1100 AA.
XX
AC AA97748;
XX
DT 06-AUG-2001 (first entry)
XX
DE Tankyrase homologue isotype 1 protein sequence.
XX
KW Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine.
XX
OS Unidentified.
XX
PN WO200130987-A2.
XX
PD 03-MAY-2001.
XX
PF 25-OCT-2000; 2000WO-US41528.
XX
PR 25-OCT-1999; 99US-0427154.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B;
XX
DR WPI: 2001-300503/31.
XX
DR N-PSDB; AAA91487.
XX
PT Novel recombinant cell cycle polypeptide, tankyrase H useful for
PT inducing or preventing cell proliferation in cells, and for diagnosing,

treating or preventing cell cycle associated disorders such as cancer

Claim 22; Fig 3; 63pp; English.

This sequence is the Tankyrase homologue isotype 1 (TaHo-1) protein of the invention. The invention also relates to the TaHo-2 protein. The TaHo proteins are useful for inducing or preventing cell proliferation in cells, and in the study or treatment of conditions mediated by the cell cycle proteins, such as to diagnose, treat or prevent cell cycle associated disorders, preferably cancer. The TaHo coding sequences are useful as hybridisation probes, in chromosome and gene mapping and in the generation of anti-sense DNA and RNA. The coding sequences are also useful for the preparation of TaHo, for generating either transgenic animals or knock out animals which, in turn, are useful in a development and screening of therapeutically useful agents, in gene therapy, as vaccine, and for construction of hybridisation probes for mapping the gene which encodes TaHo and for the genetic analysis of individuals with genetic disorders. The TaHo proteins, and their coding sequences are useful in screening assays.

SQ Sequence 1100 AA;

Query Match 100.0%; Score 5769; DB 22; Length 1100;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGRKDVVEYLQNGASVQARDGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 60
 DB 1 GFGRKDVVEYLQNGASVQARDGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 60
 QY 61 PLHEAAATGKIDVICVILLOHGAETIRNTDGTALDADPSAKAVLTGEYKKBELLESAR 120
 DB 61 PLHEAAATGKIDVICVILLOHGAETIRNTDGTALDADPSAKAVLTGEYKKBELLESAR 120
 QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHAGYNRVKIVQLLQHGADVHAKDGDLY 180
 DB 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHAGYNRVKIVQLLQHGADVHAKDGDLY 180
 QY 181 PLHNACSYGHEVTELLVHKHGCACVNDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
 DB 181 PLHNACSYGHEVTELLVHKHGCACVNDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
 QY 241 NCHNKAIDLAAPTQPKERLAYEFKHSLLQAAREADVTRIKKHSLEWMVNFKHQTHET 300
 DB 241 NCHNKAIDLAAPTQPKERLAYEFKHSLLQAAREADVTRIKKHSLEWMVNFKHQTHET 300
 QY 301 ALHCAASPYPKRKOICEILLRKGANTNEKTKEFLPLHVASEKAHNDVVVVVKEAKV 360
 DB 301 ALHCAASPYPKRKOICEILLRKGANTNEKTKEFLPLHVASEKAHNDVVVVVKEAKV 360
 QY 361 NALDNLGOTSLHRAAYCGHLOTCRLLSYGCDPNIIISLQFTALQMGNEVQOLLQEGIS 420
 DB 361 NALDNLGOTSLHRAAYCGHLOTCRLLSYGCDPNIIISLQFTALQMGNEVQOLLQEGIS 420
 QY 421 LGNSEADROLLEAAKAGDVETVKKLTQVSNCRDTEGROSTPLHFAAGYNRVSVVEYLL 480
 DB 421 LGNSEADROLLEAAKAGDVETVKKLTQVSNCRDTEGROSTPLHFAAGYNRVSVVEYLL 480
 QY 481 QHGADVHAKDKGLVPLHNACSYGHEVTELLVHKHGCACVNDLWQFTPLHEAASKNRVEV 540
 DB 481 QHGADVHAKDKGLVPLHNACSYGHEVTELLVHKHGCACVNDLWQFTPLHEAASKNRVEV 540
 QY 541 ICKLLLOHGAETPKKRDGNTPLDLVKDGTDTIQLLRGDAALLDAAKKGCARVKKLSS 600
 DB 541 ICKLLLOHGAETPKKRDGNTPLDLVKDGTDTIQLLRGDAALLDAAKKGCARVKKLSS 600
 QY 601 PDVNCRDTPQGRHSTPLHLAAGYNNEVEAYELLYLQHGADVNAODKGLIPLHNAASVGHVD 660
 DB 601 PDVNCRDTPQGRHSTPLHLAAGYNNEVEAYELLYLQHGADVNAODKGLIPLHNAASVGHVD 660
 QY 661 VAALLIKYNACVNATDKWAFPLHEAAQKGRTOQLCALLAHGADPTLKNQEGOTPLDLVS 720
 DB 661 VAALLIKYNACVNATDKWAFPLHEAAQKGRTOQLCALLAHGADPTLKNQEGOTPLDLVS 720

DB 661 VAALLIKYNACVNATDKWAFPLHEAAQKGRTOQLCALLAHGADPTLKNQEGOTPLDLVS 720
 QY 721 ADDVSALLTAAMPSPALPCYKQVNLNGVRSPGATADALSSGSPSSLSAASSLDNLSG 780
 DB 721 ADDVSALLTAAMPSPALPCYKQVNLNGVRSPGATADALSSGSPSSLSAASSLDNLSG 780
 QY 781 SFSSELSSVSSSGTEGASLSEKKEVPGVDFSIQTQVRNLGLEHLMDFEREOITLDVLVE 840
 DB 781 SFSSELSSVSSSGTEGASLSEKKEVPGVDFSIQTQVRNLGLEHLMDFEREOITLDVLVE 840
 QY 841 MGHKELKEIGNAYGHRHKLKIGVERLISGOGLNPYLTLNTSGSSTLIDLSPDKBEFQ 900
 DB 841 MGHKELKEIGNAYGHRHKLKIGVERLISGOGLNPYLTLNTSGSSTLIDLSPDKBEFQ 900
 QY 901 SVEEMOSTVREHRDGGHAGGIFNRYNLTOKVCNKKLWERYTHRRKEVSEENHNHANE 960
 DB 901 SVEEMOSTVREHRDGGHAGGIFNRYNLTOKVCNKKLWERYTHRRKEVSEENHNHANE 960
 QY 961 RMLFHGSPFVNAILHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYIGGGTGPVHKD 1020
 DB 961 RMLFHGSPFVNAILHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYIGGGTGPVHKD 1020
 QY 1021 RSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPGHHSVTGRPSVNGIALAAYVYIRGEQ 1080
 DB 1021 RSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPGHHSVTGRPSVNGIALAAYVYIRGEQ 1080
 QY 1081 AYPEYLITYQIMRPEGMVDG 1100
 DB 1081 AYPEYLITYQIMRPEGMVDG 1100

RESULT 2

AAB66295

ID AAB66295 standard; Protein; 1166 AA.

XX AAB66295;

AC AAB66295;

XX AAB66295;

DT 05-APR-2001 (first entry)

XX Human tankyrase2 TANK2-SHORT SEQ ID NO: 135.

DE Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
 KW Inflammatory disorder.

XX Homo sapiens.

OS Homo sapiens.

XX WO200100849-A1.

PN WO200100849-A1.

XX 04-JAN-2001.

PD 04-JAN-2001.

XX 28-JUN-2000; 2000WO-US17827.

PF 28-JUN-2000; 2000WO-US17827.

XX 29-JUN-1999; 99US-0141582.

PR 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

PA (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPI: 2001-102896/11.

DR N-PSDB; AAF63953.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by

PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,

PT inflammatory and autoimmune disorders -

XX Claim 3; Page 200-203; 242pp; English.

XX The present invention provides the protein and coding sequence for the

CC human tankyrase2 protein. This is found in two different versions,

CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has

CC polyADP-ribosylation activity and is involved in the modification of

CC TRF1, which is a telomere-specific binding protein. The regulation of

CC telomere length, in which TRF1 has a role, is linked to ageing and

CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
SQ Sequence 1166 AA;
Query Match 99.9%; Score 5766; DB 22; Length 1166;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGRKDVVEYLLQNGASVOARDGGLIPLHNACSGHAEVNNLLRHGADPNARDNNWYT 60
Db 67 GGRKDVVEYLLQNGANVOARDGGLIPLHNACSGHAEVNNLLRHGADPNARDNNWYT 126
QY 61 PLHEAAIKGIDVCIVLLQHGAEPTIRNTDGRATDLADPSAKAVLTGEYKKDELLESAR 120
Db 127 PLHEAAIKGIDVCIVLLQHGAEPTIRNTDGRATDLADPSAKAVLTGEYKKDELLESAR 186
QY 121 SGNEEKMMALLPLNVNCHASDGRKSTPLHLAAGYNNRVKIVOLLQHGADVHAKDKGDLV 180
Db 187 SGNEEKMMALLPLNVNCHASDGRKSTPLHLAAGYNNRVKIVOLLQHGADVHAKDKGDLV 246
QY 181 PLHNACSYGHEVTELLVYKHGACVNMADLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
Db 247 PLHNACSYGHEVTELLVYKHGACVNMADLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 306
QY 241 NCHNKSADLAPTPQKLERLAYEFKHSLLQAAAREADVTIRIKKHLISLEVMNFKHPQTHET 300
Db 307 NCHNKSADLAPTPQKLERLAYEFKHSLLQAAAREADVTIRIKKHLISLEVMNFKHPQTHET 366
QY 301 ALHCAASVPKPKOICEILLRRKANINEKTEFTPLHVASEKAHNDVVEVVKHAEKV 360
Db 367 ALHCAASVPKPKOICEILLRRKANINEKTEFTPLHVASEKAHNDVVEVVKHAEKV 426
QY 361 NALDNLGQTSLSHRAAYCGHLQTCRLLSYGCDDPNITSLQGTALQMGNEVVOQLQEGIS 420
Db 427 NALDNLGQTSLSHRAAYCGHLQTCRLLSYGCDDPNITSLQGTALQMGNEVVOQLQEGIS 486
QY 421 LGNSEADRLLEAAKAGDVETVKKLCTVQSVNCRDTEGROSTPLHEAAGYNNRVVVEYLL 480
Db 487 LGNSEADRLLEAAKAGDVETVKKLCTVQSVNCRDTEGROSTPLHEAAGYNNRVVVEYLL 546
QY 481 OHGADVHAKDKGGLVPLHNACSYGHEVAPLLVKKHGVNVDLWKFPTPLHEAAAKGKYE 540
Db 547 OHGADVHAKDKGGLVPLHNACSYGHEVAPLLVKKHGVNVDLWKFPTPLHEAAAKGKYE 606
QY 541 ICKLLQHGADPTKKNRDGNTPDLVKGDTDIQDLRGDAALLDAKKGCLARVKKLS 600
Db 607 ICKLLQHGADPTKKNRDGNTPDLVKGDTDIQDLRGDAALLDAKKGCLARVKKLS 666
QY 601 PDVNCRDTOGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAODKGLIPLHNAASYGHVD 660
Db 667 PDVNCRDTOGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAODKGLIPLHNAASYGHVD 726
QY 661 VAALLIKYNACVNATDKWAFPTPLHEAAQKRTQLCALLAHAGADPTLKNQEGTPLDLVS 720
Db 727 VAALLIKYNACVNATDKWAFPTPLHEAAQKRTQLCALLAHAGADPTLKNQEGTPLDLVS 786
QY 721 ADDYSALLTAMPSPALPSCYKPOVLNCRVSPGATADALSSGSPSSLSAASLNLG 780
Db 787 ADDYSALLTAMPSPALPSCYKPOVLNCRVSPGATADALSSGSPSSLSAASLNLG 846
QY 781 SFSSELSSVSSSGTEGASSLEKKEVPQVDFSTQFVNLGLIEHLMIDIFEREQITLDVIVE 840
Db 847 SFSSELSSVSSSGTEGASSLEKKEVPQVDFSTQFVNLGLIEHLMIDIFEREQITLDVIVE 906
QY 841 MGHEKELIGINAYGHRHKLKIGVERLIISGOOGLNPNLYTLTNTSGGTILIDLSPDKKEFQ 900
Db 907 MGHEKELIGINAYGHRHKLKIGVERLIISGOOGLNPNLYTLTNTSGGTILIDLSPDKKEFQ 966
QY 901 SVEEMQSTVREHRRDGGHAGGIFRNRYNLTKQKCNKMLWERYTHRKEVSEENHNHANE 960
Db 967 SVEEMQSTVREHRRDGGHAGGIFRNRYNLTKQKCNKMLWERYTHRKEVSEENHNHANE 1026

QY 961 RMLPHGSPFYNAIIHKGFDERHAYIGGMFCAGIYFAENSSKSNQYVYIGGGTGCPVHKD 1020
Db 1027 RMLPHGSPFYNAIIHKGFDERHAYIGGMFCAGIYFAENSSKSNQYVYIGGGTGCPVHKD 1086
QY 1021 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIRGEQ 1080
Db 1087 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIRGEQ 1146
QY 1081 AYPEYLITYQIMRPEGMDVG 1100
Db 1147 AYPEYLITYQIMRPEGMDVG 1166
RESULT 3
AAB66278
ID AAB66278 standard; Protein: 1169 AA.
XX
AC AAB66278;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 related protein sequence SEO ID NO: 2.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17827.
XX
PR 29-JUN-1999; 99US-0141582.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
DR WPI: 2001-102896/11.
XX
DR N-PSDB; AAF63837.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX
PS Disclosure: Page 109-113; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
SQ Sequence 1169 AA;
Query Match 99.9%; Score 5766; DB 22; Length 1169;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGRKDVVEYLLQNGASVOARDGGLIPLHNACSGHAEVNNLLRHGADPNARDNNWYT 60
Db 70 GGRKDVVEYLLQNGANVOARDGGLIPLHNACSGHAEVNNLLRHGADPNARDNNWYT 129
QY 61 PLHEAAIKGIDVCIVLLQHGAEPTIRNTDGRATDLADPSAKAVLTGEYKKDELLESAR 120
Db 130 PLHEAAIKGIDVCIVLLQHGAEPTIRNTDGRATDLADPSAKAVLTGEYKKDELLESAR 189
QY 121 SGNEEKMMALLPLNVNCHASDGRKSTPLHLAAGYNNRVKIVOLLQHGADVHAKDKGDLV 180

Db 190 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGDLV 249
QY 181 PLHNACSYGHEVTELLVKGACVAMDLWQFTPLHEAASKNRVEVCSSLLSYGADPTLL 240
Db 250 PLHNACSYGHEVTELLVKGACVAMDLWQFTPLHEAASKNRVEVCSSLLSYGADPTLL 309
QY 241 NCHNKSALDAPTLPOLKERLAYEFKGHSLLQAAAREADVTRIKKHLSELMVNFKHPQTHET 300
Db 310 NCHNKSALDAPTLPOLKERLAYEFKGHSLLQAAAREADVTRIKKHLSELMVNFKHPQTHET 369
QY 301 ALHCAASPYPRKQICELLRLKGANINEKTKFELTPLHVASEKAHNDVVEVVVKHEAKV 360
Db 370 ALHCAASPYPRKQICELLRLKGANINEKTKFELTPLHVASEKAHNDVVEVVVKHEAKV 429
QY 361 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLOGFTALQMGNEVQVLLQEGIS 420
Db 430 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLOGFTALQMGNEVQVLLQEGIS 489
QY 421 LGNSEADRLLEAAKAGDVETVKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
Db 490 LGNSEADRLLEAAKAGDVETVKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 549
QY 481 QHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKGAVVNVADJWKFTPLHEAAAKGYE 540
Db 550 QHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKGAVVNVADJWKFTPLHEAAAKGYE 609
QY 541 ICKLLQHGADPTKKNRDNPTPLDLVKDGDPTDIDLLRGDAALLDAKKGCLARVKKLSS 600
Db 610 ICKLLQHGADPTKKNRDNPTPLDLVKDGDPTDIDLLRGDAALLDAKKGCLARVKKLSS 669
QY 601 PDVNCRTQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVD 660
Db 670 PDVNCRTQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVD 729
QY 661 VAALLIKYNACVNATDKWAFPLHEAAKQRTQCALLLAHGADPTLKNQEGQTPDLVLS 720
Db 730 VAALLIKYNACVNATDKWAFPLHEAAKQRTQCALLLAHGADPTLKNQEGQTPDLVLS 789
QY 721 ADDVSALLTAAMPSPALPSCYKPOVLNVRSPGATADALSSGSPSSLSAASLDNLG 780
Db 790 ADDVSALLTAAMPSPALPSCYKPOVLNVRSPGATADALSSGSPSSLSAASLDNLG 849
QY 781 SFSLSWVSSSGTEGASSLEKKEVPVDFSIQFVRLNGLHLMDFEREQITDLVLE 840
Db 850 SFSLSWVSSSGTEGASSLEKKEVPVDFSIQFVRLNGLHLMDFEREQITDLVLE 909
QY 841 MGHKELKEIGINAYGHRHKLKIGVERLISGOQGLNPYLTINTSGSGTILIDLSPDDKEFO 900
Db 910 MGHKELKEIGINAYGHRHKLKIGVERLISGOQGLNPYLTINTSGSGTILIDLSPDDKEFO 969
QY 901 SVEEEMQSTVREHRDGHAGGIFNRYNLIKQVCNKKLWERYTHRRKEYSEENHNHANE 1029
Db 970 SVEEEMQSTVREHRDGHAGGIFNRYNLIKQVCNKKLWERYTHRRKEYSEENHNHANE 1099
QY 961 RMLFHGSPFFNAILHKGFDERHAYIGMFCAGIYFAENSCKSNQYVIGGGTGCPVHKD 1020
Db 1030 RMLFHGSPFFNAILHKGFDERHAYIGMFCAGIYFAENSCKSNQYVIGGGTGCPVHKD 1089
QY 1021 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHSHVTSGRPSVNGLALAEVYVIRGEQ 1080
Db 1090 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHSHVTSGRPSVNGLALAEVYVIRGEQ 1149
QY 1081 AYPEYLITQIMRPEGMVDG 1100
Db 1150 AYPEYLITQIMRPEGMVDG 1169

RESULT 4
ID AAB66288
XX AAB66288 standard; Protein; 1169 AA.
AC AAB66288;

XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone consensus protein SEQ ID NO: 101.
XX
KW Human: tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17827.
XX
PR 29-JUN-1999; 99US-0141582.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
DR WPI; 2001-102896/11.
XX
DR N-PSDB; AAF63926.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
XX inflammatory and autoimmune disorders.
PS Example 1; Page 162-1665; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
SQ Sequence 1169 AA;

Query Match 99.9%; Score 5766; DB 22; Length 1169;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGRKDVVEYLLQNGASVQARDGGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 60
Db 70 GFGRKDVVEYLLQNGANVQARDGGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 129
QY 61 PLHEAAIKGKIDYCVILQHGAEPTIRNTDGTALDADPSAKAVLTGEYKKDELLESAR 120
Db 130 PLHEAAIKGKIDYCVILQHGAEPTIRNTDGTALDADPSAKAVLTGEYKKDELLESAR 189
QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGDLV 180
Db 190 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGDLV 249
QY 181 PLHNACSYGHEVTELLVKGACVAMDLWQFTPLHEAASKNRVEVCSSLLSYGADPTLL 240
Db 250 PLHNACSYGHEVTELLVKGACVAMDLWQFTPLHEAASKNRVEVCSSLLSYGADPTLL 309
QY 241 NCHNKSALDAPTLPOLKERLAYEFKGHSLLQAAAREADVTRIKKHLSELMVNFKHPQTHET 300
Db 310 NCHNKSALDAPTLPOLKERLAYEFKGHSLLQAAAREADVTRIKKHLSELMVNFKHPQTHET 369
QY 301 ALHCAASPYPRKQICELLRLKGANINEKTKFELTPLHVASEKAHNDVVEVVVKHEAKV 360
Db 370 ALHCAASPYPRKQICELLRLKGANINEKTKFELTPLHVASEKAHNDVVEVVVKHEAKV 429
QY 361 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLOGFTALQMGNEVQVLLQEGIS 420
Db 430 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLOGFTALQMGNEVQVLLQEGIS 489

```
QY 421 LGNSEADQLLEAAKAGDVETVKKLCTVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 480
DB 490 LGNSEADQLLEAAKAGDVETVKKLCTVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 549
QY 481 QHGADVHAKDKGGLVPLHNACSYGHEVAELLVKGAVNVADLWKFTPLHEAAAKGYE 540
DB 550 QHGADVHAKDKGGLVPLHNACSYGHEVAELLVKGAVNVADLWKFTPLHEAAAKGYE 609
QY 541 ICKLLQHGADPTKKNRDGNTPLDLVKGDDTDIODLLRGGDAALLDAKKGCLARVKKLSS 600
DB 610 ICKLLQHGADPTKKNRDGNTPLDLVKGDDTDIODLLRGGDAALLDAKKGCLARVKKLSS 669
QY 601 PDNVNCRDTQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVD 660
DB 670 PDNVNCRDTQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVD 729
QY 661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTOALCALLAHAGADPTLKNQEGTDLVLS 720
DB 730 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTOALCALLAHAGADPTLKNQEGTDLVLS 789
QY 721 ADDVSALLTAAMPSPALPCYKQVNLGVRSFGATADALSSGSPSPSSLSAASLDNLGS 780
DB 790 ADDVSALLTAAMPSPALPCYKQVNLGVRSFGATADALSSGSPSPSSLSAASLDNLGS 849
QY 781 SFSELSVVSSSGTEGASLEKKEVPGVDFESTQFVRNLGLEHLMDFEREQITLDVLVE 840
DB 850 SFSELSVVSSSGTEGASLEKKEVPGVDFESTQFVRNLGLEHLMDFEREQITLDVLVE 909
QY 841 MGHKELKEIGINAYGHRHKLIGVERLISGQOGLNPYTLNTSGSTILIDLSPDDKEFQ 900
DB 910 MGHKELKEIGINAYGHRHKLIGVERLISGQOGLNPYTLNTSGSTILIDLSPDDKEFQ 969
QY 901 SVEEEMQSVRHRDGGHAGGTFNRYNLIKQKCNKKLWERYTHRRKEVSEENHNHANE 960
DB 970 SVEEEMQSVRHRDGGHAGGTFNRYNLIKQKCNKKLWERYTHRRKEVSEENHNHANE 1029
QY 961 RMLFHGSPFVNALIHKGDFERHAYIGMGFAGIYFAENSCKSNQYVYIGGTCGCPVHKD 1020
DB 1030 RMLFHGSPFVNALIHKGDFERHAYIGMGFAGIYFAENSCKSNQYVYIGGTCGCPVHKD 1089
QY 1021 RSCYICHRQLLCFVRTLGKSFQFSAMKMAHSPPGHSHSVTGRPSVNGLALAEYVIYRGEQ 1080
DB 1090 RSCYICHRQLLCFVRTLGKSFQFSAMKMAHSPPGHSHSVTGRPSVNGLALAEYVIYRGEQ 1149
QY 1081 AYPEYLITYOIMRPEGWVDG 1100
DB 1150 AYPEYLITYOIMRPEGWVDG 1169

RESULT 5
ID AAY97749
XX AAY97749 standard; Protein; 1240 AA.
AC AAY97749;
XX AAY97749;
DT 06-AUG-2001 (first entry)
XX Tankyrase homologue isotype 2 protein sequence.
DE Tankyrase homologue isotype 2; TaHo-1; TaHo-2; cell proliferation;
XX cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine.
XX Unidentified.
OS Unidentified.
XX W0200130987-A2.
XX 03-MAY-2001.
XX 25-OCT-2000; 2000WO-US41528.
XX 25-OCT-1999; 99US-0427154.
PR
```

```
XX (RIGE-) RIGEL PHARM INC.
PA Luo Y, Chan E, Xu X, Huang B;
PI WPI; 2001-300503/31.
DR N-PSDB; AAA91488.
XX Novel recombinant cell cycle polypeptide, tankyrase H useful for
PT inducing or preventing cell proliferation in cells, and for diagnosing,
PT treating or preventing cell cycle associated disorders such as cancer
PT
XX Claim 22; Fig 4; 63pp; English.
PS This sequence is the Tankyrase homologue isotype 2 (TaHo-2) protein
XX of the invention. The invention also relates to the TaHo-2 protein.
CC The TaHo proteins are useful for inducing or preventing cell
CC proliferation in cells, and in the study or treatment of conditions
CC mediated by the cell cycle proteins, such as to diagnose, treat or
CC prevent cell cycle associated disorders, preferably cancer. The TaHo
CC coding sequences are useful as hybridisation probes, in chromosome and
CC gene mapping and in the generation of anti-sense DNA and RNA. The coding
CC sequences are also useful for the preparation of TaHo, for generating
CC either transgenic animals or knock out animals which, in turn, are useful
CC in a development and screening of therapeutically useful agents, in gene
CC therapy, as vaccine, and for construction of hybridisation probes for
CC mapping the gene which encodes TaHo and for the genetic analysis of
CC individuals with genetic disorders. The TaHo proteins, and their coding
CC sequences are useful in screening assays.
XX
SQ Sequence 1240 AA;
Query Match 99.9%; Score 5766; DB 22; Length 1240;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GFGRKDVVEYLLONGASVQARDGGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNYNT 60
DB 141 GFGRKDVVEYLLONGANVQARDGGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNYNT 200
QY 61 PLHEAAIKGKIDVCIIVLLOHGAETPIRNTDGTALDLPADPSAKAVLTGEYKKDELLESAR 120
DB 201 PLHEAAIKGKIDVCIIVLLOHGAETPIRNTDGTALDLPADPSAKAVLTGEYKKDELLESAR 260
QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGDLV 180
DB 261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGDLV 320
QY 181 PLHNACSYGHEVTELLVKGACVNAEDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
DB 321 PLHNACSYGHEVTELLVKGACVNAEDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
QY 241 NCHNKSADLAPTPOKERLAYEFKSHLSLQAAAREADVTRIKKHLSEMVNFKHPQTHET 300
DB 381 NCHNKSADLAPTPOKERLAYEFKSHLSLQAAAREADVTRIKKHLSEMVNFKHPQTHET 440
QY 301 ALHCAASAPYKPKKOICEILLRKGANINEKTEFTPLHVASEKAHNDVVEVYVYKHEAV 360
DB 441 ALHCAASAPYKPKKOICEILLRKGANINEKTEFTPLHVASEKAHNDVVEVYVYKHEAV 500
QY 361 NALDNLGOTSLHRAAYCGHLQTCRLLLSVCGDPNIIISLQGTALOMGNENVOQLLOEGIS 420
DB 501 NALDNLGOTSLHRAAYCGHLQTCRLLLSVCGDPNIIISLQGTALOMGNENVOQLLOEGIS 560
QY 421 LGNSEADROLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
DB 561 LGNSEADROLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620
QY 481 QHGADVHAKDKGGLVPLHNACSYGHEVAELLVKGAVNVADLWKFTPLHEAAAKGYE 540
DB 621 QHGADVHAKDKGGLVPLHNACSYGHEVAELLVKGAVNVADLWKFTPLHEAAAKGYE 680
```

QY 541 ICKLLQHGADPTKKNRDNPTPLDLVKDGDITDQLLRGDAALLDAKKGCLARVKKLSS 600
 Db 681 ICKLLQHGADPTKKNRDNPTPLDLVKDGDITDQLLRGDAALLDAKKGCLARVKKLSS 740
 QY 601 PDVNCRDTOGRHSPTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVD 660
 Db 741 PDVNCRDTOGRHSPTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVD 800
 QY 661 VAALLIKYNACVNATDKWAFPLHEAAQKGRITQCALLLAHGADPTLKNQSGQPTDLVS 720
 Db 801 VAALLIKYNACVNATDKWAFPLHEAAQKGRITQCALLLAHGADPTLKNQSGQPTDLVS 860
 QY 721 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASLDNLG 780
 Db 861 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASLDNLG 920
 QY 781 SFSLSVSSVSSGTEGASSLEKKEVPVDFSTITQFVRNLGLEHLMDFEREQITDLVLE 840
 Db 921 SFSLSVSSVSSGTEGASSLEKKEVPVDFSTITQFVRNLGLEHLMDFEREQITDLVLE 980
 QY 841 MGHKELKEIGINAYGHRHKLKIGVERLISGQGLNPYLTNTSGSGTILIDLSPDDKEFQ 900
 Db 981 MGHKELKEIGINAYGHRHKLKIGVERLISGQGLNPYLTNTSGSGTILIDLSPDDKEFQ 1040
 QY 901 SVREEMOSTVREHHDGAGGIFENRYNLIKQVCNKKLWERTHRKKEVSEENHNHANE 960
 Db 1041 SVREEMOSTVREHHDGAGGIFENRYNLIKQVCNKKLWERTHRKKEVSEENHNHANE 1100
 QY 961 RMLFHGSPFFNAIHKGFDERHAYIGMFCAGIYFAENSSKSNQYVYIGGGTGPVHKD 1020
 Db 1101 RMLFHGSPFFNAIHKGFDERHAYIGMFCAGIYFAENSSKSNQYVYIGGGTGPVHKD 1160
 QY 1021 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSHVTRPSVNGLALAEYVYIRGEQ 1080
 Db 1161 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSHVTRPSVNGLALAEYVYIRGEQ 1220
 QY 1081 AYPEYLITYQIMRPEGMDVG 1100
 Db 1221 AYPEYLITYQIMRPEGMDVG 1240

RESULT 6
 AAB66290
 ID AAB66290 standard; Protein; 1262 AA.
 AC AAB66290;
 XX
 DT 05-APR-2001 (first entry)
 XX
 DE Human tankyrase2 clone consensus protein SEQ ID NO: 107.
 XX
 KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
 KW inflammatory disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200100849-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 28-JUN-2000; 2000WO-US17827.
 XX
 PR 29-JUN-1999; 99US-0141582.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
 XX
 DR WPI; 2001-102896/11.
 XX
 DR N-PSDB; AAF63930.
 XX
 PT New tankyrase2 polypeptides, useful for treating conditions mediated by
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,

PT inflammatory and autoimmune disorders -
 XX
 PS Example 2; Page 173-176; 242pp; English.

XX The present invention provides the protein and coding sequence for the
 CC human tankyrase2 protein. This is found in two different versions,
 CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
 CC polyADP-ribosylation activity and is involved in the modification of
 CC TRF1, which is a telomere-specific binding protein. The regulation of
 CC telomere length, in which TRF1 has a role, is linked to ageing and
 CC cancer. The sequences are useful in the treatment of cancers and
 CC inflammatory disorders.

XX Sequence 1262 AA;

Query Match 99.9%; Score 5766; DB 22; Length 1262;
 Best local Similarity 99.9%; Pred. No. 0;
 Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGKQDVVEYLLQNGASVQARDGGLPLHNACSGFHAENVNLLLRHGADPNARDNNYNT 60
 Db 163 GFGKRDVVEYLLQNGANVQARDGGLPLHNACSGFHAENVNLLLRHGADPNARDNNYNT 222
 QY 61 PLHEAALIKGIDVCIVLLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKKEDELLESAR 120
 Db 223 PLHEAALIKGIDVCIVLLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKKEDELLESAR 282
 QY 121 SGNEEKMAALLPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGDLV 180
 Db 283 SGNEEKMAALLPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGDLV 342
 QY 181 PLHNACSYGHEYTELLVKGACVNMALWQFTPLHEAASKNRVEVCSLLSYGADPTLL 240
 Db 343 PLHNACSYGHEYTELLVKGACVNMALWQFTPLHEAASKNRVEVCSLLSYGADPTLL 402
 QY 241 NCHNKSAIDLAPTQPKERLAYEFKGHSLLQAAAREADVTRIKKHSLEMYNFKHPQTHET 300
 Db 403 NCHNKSAIDLAPTQPKERLAYEFKGHSLLQAAAREADVTRIKKHSLEMYNFKHPQTHET 462
 QY 301 ALHCAAAAPYKPKKQICELLRKGANINEKTEFLPLHVASEKAHNDVYVVKHEAV 360
 Db 463 ALHCAAAAPYKPKKQICELLRKGANINEKTEFLPLHVASEKAHNDVYVVKHEAV 522
 QY 361 NALDNLGQTSLHRAAYCGHLOTCTRLLSYGCDPNIIISLOGFTALQMGNEVQQLQEGIS 420
 Db 523 NALDNLGQTSLHRAAYCGHLOTCTRLLSYGCDPNIIISLOGFTALQMGNEVQQLQEGIS 582
 QY 421 LGNSEADROLLEAAKAGDVETVKKCTVQSVNCRDIEGRQSTPLHFAAGYNRYSVVEYLL 480
 Db 583 LGNSEADROLLEAAKAGDVETVKKCTVQSVNCRDIEGRQSTPLHFAAGYNRYSVVEYLL 642
 QY 481 QHGADVHAKDKGGLVPLHNACSYGHEYTELLVKGACVNMALWQFTPLHEAASKNRVEV 540
 Db 643 QHGADVHAKDKGGLVPLHNACSYGHEYTELLVKGACVNMALWQFTPLHEAASKNRVEV 702
 QY 541 ICKLLQHGADPTKKNRDNPTPLDLVKDGDITDQLLRGDAALLDAKKGCLARVKKLSS 600
 Db 703 ICKLLQHGADPTKKNRDNPTPLDLVKDGDITDQLLRGDAALLDAKKGCLARVKKLSS 762
 QY 601 PDVNCRDTOGRHSPTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVD 660
 Db 763 PDVNCRDTOGRHSPTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVD 822
 QY 661 VAALLIKYNACVNATDKWAFPLHEAAQKGRITQCALLLAHGADPTLKNQSGQPTDLVS 720
 Db 823 VAALLIKYNACVNATDKWAFPLHEAAQKGRITQCALLLAHGADPTLKNQSGQPTDLVS 882
 QY 721 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASLDNLG 780
 Db 883 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASLDNLG 942
 QY 781 SFSLSVSSVSSGTEGASSLEKKEVPVDFSTITQFVRNLGLEHLMDFEREQITDLVLE 840
 Db 942 SFSLSVSSVSSGTEGASSLEKKEVPVDFSTITQFVRNLGLEHLMDFEREQITDLVLE 1000

Db 943 SFSELSSVSSSGTEGASSLEKKEVPVDFSTQFVRNLGLEHLMDFEREOITLDVLVE 1002
Qy 841 MGHKELKEIGINAYGHRHKLKIGVERLISGOOGLNPLYLTNTSGSGTILIDLSPDDKEFQ 900
Db 1003 MGHKELKEIGINAYGHRHKLKIGVERLISGOOGLNPLYLTNTSGSGTILIDLSPDDKEFQ 1062
Qy 901 SVEEEMOSTVREHRDGGHAGGIFNRYNLIKQKVCNKKLWERYTHRRKEVSEENHNHANE 960
Db 1063 SVEEEMOSTVREHRDGGHAGGIFNRYNLIKQKVCNKKLWERYTHRRKEVSEENHNHANE 1122
Qy 961 RMLFHGSPFVNAIHKGFDERHAYIGMFGAGIYFAENSCKSNQYVYGGGTGCPVHKD 1020
Db 1123 RMLFHGSPFVNAIHKGFDERHAYIGMFGAGIYFAENSCKSNQYVYGGGTGCPVHKD 1182
Qy 1021 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYVYRGEQ 1080
Db 1183 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYVYRGEQ 1242
Qy 1081 AYPEYLITYQIMRPEGMVDG 1100
Db 1243 AYPEYLITYQIMRPEGMVDG 1262

RESULT 7
AAB66294 standard; Protein; 1385 AA.

XX AC AAB66294;
XX 05-APR-2001 (first entry)
XX Human tankyrase2 TANK2-LONG SEQ ID NO: 133.
DE Human tankyrase2 TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX Homo sapiens.
XX WO200100849-A1.
XX PD 04-JAN-2001.
XX 28-JUN-2000; 2000WO-US17827.
XX PF 29-JUN-1999; 99US-0141582.
XX PR (ICOS-) ICOS CORP.
XX PA Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX PI WPI: 2001-102896/11.
XX DR N-PSDB; AAF63952.
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX Claim 2: Page 191-194; 242pp; English.

XX The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC poly(ADP-ribose)ylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.

XX Sequence 1385 AA:

Query Match 99.9%; Score 5766; DB 22; Length 1385;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFGKRDVVEYLLONGCASVOARDGGLIPLHNACSGHAEVNVNLLRHGADPNARDNNWYT 60
Db 286 GFGKRDVVEYLLONGANVOARDGGLIPLHNACSGHAEVNVNLLRHGADPNARDNNWYT 345
Qy 61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRALDLADPSAKAVLTGEYKDELLESAR 120
Db 346 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRALDLADPSAKAVLTGEYKDELLESAR 405
Qy 121 SNEEKEMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLOHGADVHAKDKGLV 180
Db 406 SNEEKEMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLOHGADVHAKDKGLV 465
Qy 181 PLHNACSYGHYEVTELLVYKHGACVNAWDLWQFTPLHEAASKNRVEVCSSLISYGAADPTLL 240
Db 466 PLHNACSYGHYEVTELLVYKHGACVNAWDLWQFTPLHEAASKNRVEVCSSLISYGAADPTLL 525
Qy 241 NCHNKSALDAPTPQIKERLAYEFKGHSLLQAARADVTIRIKKHLSEWVNFKHPOTHE 300
Db 526 NCHNKSALDAPTPQIKERLAYEFKGHSLLQAARADVTIRIKKHLSEWVNFKHPOTHE 585
Qy 301 ALHCAAAAPYPRKKQICELLRLKGANINEKTEFELTPLHVASEKAHNDVVEVVKHEAKV 360
Db 586 ALHCAAAAPYPRKKQICELLRLKGANINEKTEFELTPLHVASEKAHNDVVEVVKHEAKV 645
Qy 361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIISLQFTALQMGNEVVOQLQEGIS 420
Db 646 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIISLQFTALQMGNEVVOQLQEGIS 705
Qy 421 LGNSEADROLLEAAKAGDVETVKKLTQVSNVCRDTEGRQSTPLHFAAGYNRVSVVEYLL 480
Db 706 LGNSEADROLLEAAKAGDVETVKKLTQVSNVCRDTEGRQSTPLHFAAGYNRVSVVEYLL 765
Qy 481 QHGADVHAKDKGGLVPLHNACSYGHYEAELLVKHGAVNVADLWKFETPLHEAAAKGKYE 540
Db 766 QHGADVHAKDKGGLVPLHNACSYGHYEAELLVKHGAVNVADLWKFETPLHEAAAKGKYE 825
Qy 541 ICKLLLOHGADPTKKNRDGNTPLDLVKDGDTDIQLLGRDAALLDAKKGCLARVKKLS 600
Db 826 ICKLLLOHGADPTKKNRDGNTPLDLVKDGDTDIQLLGRDAALLDAKKGCLARVKKLS 885
Qy 601 PDVNCRDGTGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVD 660
Db 886 PDVNCRDGTGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVD 945
Qy 661 VAALLIKYNACVNATDKWAFETPLHEAAQKGRTOCALLLAHGAADPTLKNOEGOTPLDLVS 720
Db 946 VAALLIKYNACVNATDKWAFETPLHEAAQKGRTOCALLLAHGAADPTLKNOEGOTPLDLVS 1005
Qy 721 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASLDNLISG 780
Db 1006 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASLDNLISG 1065
Qy 781 SFSSELSSVSSSGTEGASSLEKKEVPVDFSTQFVRNLGLEHLMDFEREOITLDVLVE 840
Db 1066 SFSSELSSVSSSGTEGASSLEKKEVPVDFSTQFVRNLGLEHLMDFEREOITLDVLVE 1125
Qy 841 MGHKELKEIGINAYGHRHKLKIGVERLISGOOGLNPLYLTNTSGSGTILIDLSPDDKEFQ 900
Db 1126 MGHKELKEIGINAYGHRHKLKIGVERLISGOOGLNPLYLTNTSGSGTILIDLSPDDKEFQ 1185
Qy 901 SVEEEMOSTVREHRDGGHAGGIFNRYNLIKQKVCNKKLWERYTHRRKEVSEENHNHANE 960
Db 1186 SVEEEMOSTVREHRDGGHAGGIFNRYNLIKQKVCNKKLWERYTHRRKEVSEENHNHANE 1245
Qy 961 RMLFHGSPFVNAIHKGFDERHAYIGMFGAGIYFAENSCKSNQYVYGGGTGCPVHKD 1020
Db 1246 RMLFHGSPFVNAIHKGFDERHAYIGMFGAGIYFAENSCKSNQYVYGGGTGCPVHKD 1305
Qy 1021 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYVYRGEQ 1080
Db 1306 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYVYRGEQ 1365

QY 1081 AYPEYLITYQIMRPEGMDG 1100
DB 1366 AYPEYLITYQIMRPEGMDG 1385

RESULT 8
ID AAB27211 standard; Protein; 1166 AA.
XX AAB27211;
XX 27-FEB-2001 (first entry)
XX Human tankyrase II protein sequence SEQ ID NO: 6.
XX Human; tankyrase II; telomere length; signal transduction.
XX Homo sapiens.

Key Location/Qualifiers
FT Misc-difference 124 /note= "encoded by TTA"
FT Misc-difference 125 /note= "encoded by TAC"
XX WO200061813-A1.

XX 19-OCT-2000.
XX 10-APR-2000; 2000WO-US09558.
XX 09-APR-1999; 99US-0128577.
XX 13-APR-1999; 99US-0129123.

XX (GERO-) GERON CORP.
XX Morin GB, Funk WD, Piatyszek MA;
XX WPI; 2000-679503/66.
XX N-PSDB; AAC66825.

XX Novel mammalian Tankyrase II polypeptide and the polynucleotide
XX encoding the polypeptide useful for modulating or maintaining telomere
XX length, replicative capacity, apoptosis, chromosome packing or gene
XX expression
XX Claim 4; Fig 4; 52pp; English.

XX The present sequence is a version of the human tankyrase II protein
XX sequence. The protein is thought to be involved in signal transduction in
XX the cell, and to have binding activity for other telomere-associated
XX proteins. It is possible that it plays a role in the regulation of cell
XX telomere length, thus affecting the replicative ability of the cell. The
XX protein is useful for ribosylating target proteins, for determining
XX tankyrase II binding activity in a sample, and for modulating telomere
XX length in a cell.
XX Sequence 1166 AA;

Query Match 99.7%; Score 5750; DB 21; Length 1166;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1097; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GFGKRDVVEYLLQNGASVOARDGGLIPLHNACSGHAEVNVNLLRHGADPNARDNWTY 60
DB 67 GFGKRDVVEYLLQNGASVOARDGGLIPLHNACSGHAEVNVNLLRHGADPNARDNWTY 126
QY 61 PLHEAAIKGIDVCIIVLLQHGAEPTIRNTDGTALDADPSAKAVLTGTYKKDELESAR 120
DB 127 PLHEAAIKGIDVCIIVLLQHGAEPTIRNTDGTALDADPSAKAVLTGTYKKDELESAR 186
QY 121 SGNEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRYKIVOLLQHGADVHAKDGLV 180

DB 187 SGNEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRYKIVOLLQHGADVHAKDGLV 246
QY 181 PLHNACSYGHEYETELLVYKHGACVNAWDLWOFTPLHEAASKNRVVCISLLSYGADPTLL 240
DB 247 PLHNACSYGHEYETELLVYKHGACVNAWDLWOFTPLHEAASKNRVVCISLLSYGADPTLL 306
QY 241 NCHNKSATDLAPTQLKERLAYEFKSHLSLQAAREADVTIRIKKHSLEMVNFKHPQTHET 300
DB 307 NCHNKSATDLAPTQLKERLAYEFKSHLSLQAAREADVTIRIKKHSLEMVNFKHPQTHET 366
QY 301 ALHCAASPYPKRKOICELLRKGANINEKTEFTPLHVAASEKAHNDVVEVYKHEAKV 360
DB 367 ALHCAASPYPKRKOICELLRKGANINEKTEFTPLHVAASEKAHNDVVEVYKHEAKV 426
QY 361 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDPNIISLQGTALQMGNNVQOLLQEGIS 420
DB 427 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDPNIISLQGTALQMGNNVQOLLQEGIS 486
QY 421 LGNSEADROLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
DB 487 LGNSEADROLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546
QY 481 QHGADVHAKDKGGLVPLHNACSYGHEYVAELLVYKHGAVNVADLWKFTPLHEAAKGYE 540
DB 547 QHGADVHAKDKGGLVPLHNACSYGHEYVAELLVYKHGAVNVADLWKFTPLHEAAKGYE 606
QY 541 ICKLLQHGADPTKKNRDGNTPDLVKDGTDIQDLRGDAALDAKKGCLARVKKLSS 600
DB 607 ICKLLQHGADPTKKNRDGNTPDLVKDGTDIQDLRGDAALDAKKGCLARVKKLSS 666
QY 601 PDVNCRDTOGRHSTPLHLAAGYNNEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
DB 667 PDVNCRDTOGRHSTPLHLAAGYNNEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 726
QY 661 VAALLIKYNACVNATDKWFTPLHEAAKGTQALCALLLAHAGADPTLKNBEGOTPLDLVS 720
DB 727 VAALLIKYNACVNATDKWFTPLHEAAKGTQALCALLLAHAGADPTLKNBEGOTPLDLVS 786
QY 721 ADDVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSSGSPSSSLSAASLNLGSG 780
DB 787 ADDVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSSGSPSSSLSAASLNLGSG 846
QY 781 SFSELSVSVSSSGTEGASSLEKKEVPGVDFSIQFVRNLGLEHMDIFEREQITLDVIVE 840
DB 847 SFSELSVSVSSSGTEGASSLEKKEVPGVDFSIQFVRNLGLEHMDIFEREQITLDVIVE 906
QY 841 MGHKELKEIGINAYGHRHKLKIGVERLISGOQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
DB 907 MGHKELKEIGINAYGHRHKLKIGVERLISGOQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 966
QY 901 SYVEEEMQSTVREHRDGHAGGIFNRYNLIKIQVCNKKLWERYTHRRKEVSEENHNHANE 960
DB 967 SYVEEEMQSTVREHRDGHAGGIFNRYNLIKIQVCNKKLWERYTHRRKEVSEENHNHANE 1026
QY 961 RMLFHGSPFVNAILHKGDFERHAYIGGMFGAGIYFAENSCKSNQYVYGGTGCPVHKD 1020
DB 1027 RMLFHGSPFVNAILHKGDFERHAYIGGMFGAGIYFAENSCKSNQYVYGGTGCPVHKD 1086
QY 1021 RSCYICHHOLLEFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLAELAEVYVIRGEO 1080
DB 1087 RSCYICHHOLLEFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLAELAEVYVIRGEO 1146
QY 1081 AYPEYLITYQIMRPEGMDG 1100
DB 1147 AYPEYLITYQIMRPEGMDG 1166

RESULT 9
AAB272589
ID AAY72589 standard; Protein; 1166 AA.
XX AAY72589;
XX

DT 02-MAY-2001 (first entry)
XX Human tankyrase homolog protein (THP).

DE
XX
KW Human; tankyrase homolog protein; THP; gene therapy; cancer;
KW tumour; basal cell carcinoma; therapy; genetic mapping;
KW cytosstatic.

XX
OS Homo sapiens.

XX
XX Key Location/Qualifiers
FH Misc-difference 368 /label= Met, Leu, Val
FT Misc-difference 392 /label= Asn, Thr
FT Misc-difference 415 /label= Val, Ile
FT Misc-difference 558 /label= Gly, Glu
FT Misc-difference 559 /label= Gly, Ala
FT Misc-difference 764 /label= Leu, Val
FT Misc-difference 884 /label= Asn, His, Asp, Tyr

XX WO200104326-A1.
XX 18-JAN-2001.
XX 03-JUL-2000; 2000WO-EP06609.
XX 09-JUL-1999; 99US-0350982.
XX (PHAA) PHARMACIA & UPJOHN SPA.
XX Berthelsen J, Toma S, Isacchi A;
XX WPI; 2001-168422/17.
XX N-PSDB; AAD02578.

XX New tankyrase homolog protein (THP) polynucleotide and polypeptide
XX useful in gene therapy, diagnosis and treatment or prevention of
XX unregulated cell growth, such as cancer or tumor cell growth -
XX Claim 20; Page 50-54; 60pp; English.

XX The present sequence is human tankyrase homolog
XX protein (THP). The THP polypeptides and polynucleotides of the
XX invention are useful in gene therapy and for treating or preventing
XX unregulated cell growth such as cancer or tumour (e.g. basal cell
XX carcinoma). The nucleic acid molecules of the invention and their
XX fragments are useful for restriction fragment length polymorphism
XX (RFLP) associated with certain disorders, as well as for genetic
XX mapping. Antisense oligonucleotides, or fragments of nucleic acid
XX encoding THP are useful as diagnostic tools for probing the
XX expression of THP gene in various tissues. THP can be used as
XX antigens for raising antibodies against them and in assays for
XX identifying compounds that modulate their activity. They are used in
XX the manufacture of a medicament directed towards cancers or tumours.
XX THP are also useful for screening compounds in a variety of drugs
XX screening techniques and as a research tool for identification,
XX characterisation and purification of interacting, regulatory proteins.

XX Sequence 1166 AA:
Query Match 99.1%; Score 5717; DB 22; Length 1166;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1091; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
1 GFGKQVVEYLLQNGASVQARDGGLIPLHNACSFSGHAEVNNLLRHGADPNARDNNVT 60
67 GFGKQVVEYLLQNGANVQARDGGLIPLHNACSFSGHAEVNNLLRHGADPNARDNNVT 126

Qy 61 PLHEAAIKGKIDVCIVLLOHGAETPIRNTDGRALTDLADPSAKAVLTGEYKKDELLESAR 120
Db 127 PLHEAAIKGKIDVCIVLLOHGAETPIRNTDGRALTDLADPSAKAVLTGEYKKDELLESAR 186
Qy 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLV 180
Db 187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLV 246
Qy 181 PLHNACSYGHEVTELLVKHGACYNAMDLMQFTPLHEAASKNRVEVCSLLISYGADPTLL 240
Db 247 PLHNACSYGHEVTELLVKHGACYNAMDLMQFTPLHEAASKNRVEVCSLLISYGADPTLL 306
Qy 241 NCHNKSADLAPTQPKERLAYEFKGHSLLOAAREADYTRIKKHLSLEMFNFKHPQTHET 300
Db 307 NCHNKSADLAPTQPKERLAYEFKGHSLLOAAREADYTRIKKHLSLEMFNFKHPQTHET 366
Qy 301 ALHCAASPYPRKQICELELLRKGANINEKTEFTPLHVASEKAHNDVVEVVKHEAV 360
Db 367 AXHCAASPYPRKQICELELLRKGAXINEKTEFTPLHVASEKAHNDVVEVVKHEAV 426
Qy 361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIISIQFTALQMGNEVQOLLQEGIS 420
Db 427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIISIQFTALQMGNEVQOLLQEGIS 486
Qy 421 LGNSEADROLLEAAKAGDVETVKKLCVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 480
Db 487 LGNSEADROLLEAAKAGDVETVKKLCVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 546
Qy 481 QHGADVHAKDKGLVPLHNACSYGHEVAEELVKHGAVNVADLWKFTPLHEAANKGYE 540
Db 547 QHGADVHAKDKXXLVLPLHNACSYGHEVAEELVKHGAVNVADLWKFTPLHEAANKGYE 606
Qy 541 ICKLLQHGADPTKKNRDGNTPLDLVKDGTDIQDLLRGDAALLDAKKGCLARVKLSS 600
Db 607 ICKLLQHGADPTKKNRDGNTPLDLVKDGTDIQDLLRGDAALLDAKKGCLARVKLSS 666
Qy 601 PDVNCRDQTGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
Db 667 PDVNCRDQTGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 726
Qy 661 VAALLIKYNACVNATDKWFTPLHEAAKQKRTOLCALLAHGADPTLKNQEGOTPLDLVS 720
Db 727 VAALLIKYNACVNATDKWFTPLHEAAKQKRTOLCALLAHGADPTLKNQEGOTPLDLVS 786
Qy 721 ADDVSALLTAAMPSPALPCYKQVLNVRSPGATADALSSGSPSSLSAASLDNLSC 780
Db 787 ADDVSALLTAAMPSPALPCYKQVLNVRSPGATADALSSGSPSSLSAASLDNLSC 846
Qy 781 SFSELSSVSSSGTEGASSLEKKEVPGVDSITQFVRNLGLEHLMDFEREQITLDLVE 840
Db 847 SFSELSSVSSSGTEGASSLEKKEVPGVDSITQFVRNLGLEHLMDFEREQITLDLVE 906
Qy 841 MGHKELKEIGNAYGHRHKLKVERLISGQGLNPYLTNTSGSGTILIDLSPPDKKEFQ 900
Db 907 MGHKELKEIGNAYGHRHKLKVERLISGQGLNPYLTNTSGSGTILIDLSPPDKKEFQ 966
Qy 901 SVEEMQSTVREHRDGGHAGGIFNRYNLTKQVCNKKLMERYTHRRKEVSEENHNHANE 960
Db 967 SVEEMQSTVREHRDGGHAGGIFNRYNLTKQVCNKKLMERYTHRRKEVSEENHNHANE 1026
Qy 961 RMLFHGSPFVNAITHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGTCGPVHKD 1020
Db 1027 RMLFHGSPFVNAITHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGTCGPVHKD 1086
Qy 1021 RSCYICHRQLLFCRVTLGKSFLOFSAMKWAHSPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
Db 1087 RSCYICHRQLLFCRVTLGKSFLOFSAMKWAHSPGHHSVTGRPSVNGLALAEYVIYRGEQ 1146
Qy 1081 AYPEYLITYQIMRPEGMVDG 1100
Db 1147 AYPEYLITYQIMRPEGMVDG 1166

RESULT 10
 RAY05734
 ID AAY05734 standard; Protein; 1074 AA.
 XX
 AC AAY05734;
 XX
 DT 19-JUL-1999 (first entry)
 XX
 DE Human Grb7 effector 2.2412 protein.
 XX
 KW Grb7 effector; 2.2412 protein; human; signal transduction;
 KW tumour marker; breast cancer; prostate cancer; prognosis;
 KW diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN W09915647-A1.
 XX
 PD 01-APR-1999.
 XX
 PF 23-SEP-1998; 98WO-AU00795.
 XX
 PR 23-SEP-1997; 97AU-0009388.
 XX
 PA (GARV-) GARVAN INST MEDICAL RES.
 XX
 PI Daly RJ, Sutherland RL;
 XX
 DR WPI; 1999-254707/21.
 XX
 DR N-PSDB; AAX25366.
 XX
 PT New candidate effector for the Grb7 family of signalling proteins,
 PT and specific antibody, useful for detection and treatment of cancer
 XX
 PS Claim 9; Fig 1; 24pp; English.
 XX
 CC The present sequence represents a novel candidate effector for
 CC the Grb7 family of signalling proteins, termed 2.2412. The
 CC sequence is predicted from a partial cDNA (see AAX25366). 2.2412
 CC mRNA transcripts (7 kb) were detected in all tissues examined
 CC with the exception of kidney. Expression was particularly high in
 CC skeletal muscle and placenta. Analysis of the sequence revealed
 CC significant homology to a large number of proteins containing
 CC ankyrin-like repeats. The 2.2412 gene was localised to between
 CC chromosome 10q23.2 and proximal 10q23.32. Deletions in the
 CC 10q22-25 region have been detected in human breast, prostate,
 CC renal, small cell lung and endometrial carcinomas, glioblastoma
 CC multiforme, melanoma and meningiomas. Detection of the protein
 CC encoded by the 2.2412 cDNA in a sample should provide a useful
 CC tumour marker and/or prognostic indicator for certain human
 CC cancers, in particular breast cancer and prostate cancer.
 CC Antagonism of the interaction between Grb7 family members and the
 CC encoded protein should provide a novel treatment strategy for human
 CC diseases exhibiting aberrant receptor tyrosine kinase signalling,
 CC such as cancer. Anti-Grb7 antibodies can be used in methods of
 CC detecting the presence of 2.2412 protein in a sample.
 XX
 SQ Sequence 1074 AA;

Query Match 97.2%; Score 5605; DB 20; Length 1074;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1069; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 27 IPLHACSGFHAENVNLLRHGADPNARDNNWYTPLHAAIKGKIDVCIVLQHGAEPTI 86
 DB 1 IPLHACSGFHAENVNLLRHGADPNARDNNWYTPLHAAIKGKIDVCIVLQHGAEPTI 60
 QY 87 RNTDGTALDLADPSAKAVLTGEYKKDELLESARSGNEKMMALLTPLNVNCHASDGRKS 146
 DB 61 RNTDGTALDLADPSAKAVLTGEYKKDELLESARSGNEKMMALLTPLNVNCHASDGRKS 120
 QY 147 TPLHLAGYNNRVKIVQLLQHGADVHAKDKGLVPLHNAACSYGHYEVTELLYKKGACVNA 206

DB 121 TPLHLAGYNNRVKIVQLLQHGADVHAKDKGLVPLHNAACSYGHYEVTELLYKKGACVNA 180
 QY 207 MDLWQFTPLHAAAKNRVEVCSLLLSYGADPTLLCHNKNKSAIDLAPTQPKERLAYEFGK 266
 DB 181 MDLWQFTPLHAAAKNRVEVCSLLLSYGADPTLLCHNKNKSAIDLAPTQPKERLAYEFGK 240
 QY 267 HSLQAAAREADVTRIKKHLSLSEMVNFKHPQTHETALHCAAAAPYKPKQICELLRLKGAN 326
 DB 241 HSLQAAAREADVTRIKKHLSLSEMVNFKHPQTHETALHCAAAAPYKPKQICELLRLKGAN 300
 QY 327 INEKTKEFTPLHVASSEKANDVVEVYKHEAKVNALNGLGOTSLHRAAYCGHLOTCLRL 386
 DB 301 INEKTKEFTPLHVASSEKANDVVEVYKHEAKVNALNGLGOTSLHRAAYCGHLOTCLRL 360
 QY 387 LSYGCDPNIIISLOGFTALQMGNEVQQLQEGISLGNSEADROLLEAAKAGDVETVKKLC 446
 DB 361 LSYGCDPNIIISLOGFTALQMGNEVQQLQEGISLGNSEADROLLEAAKAGDVETVKKLC 420
 QY 447 TVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGGLVPLHNAACSYGHY 506
 DB 421 TVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGGLVPLHNAACSYGHY 480
 QY 507 EVAELLVKHGAVNVADLWKFTPLHAAAKGKYETCKILLLOHGADPTKKNRDNTPDLV 566
 DB 481 EVAELLVKHGAVNVADLWKFTPLHAAAKGKYETCKILLLOHGADPTKKNRDNTPDLV 540
 QY 567 KDGTDIQLDLRGAALLDAAKKGLARVKLSSPDNVNCRDTQGRHSTPLHLAGYNNL 626
 DB 541 KDGTDIQLDLRGAALLDAAKKGLARVKLSSPDNVNCRDTQGRHSTPLHLAGYNNL 600
 QY 627 EVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFPLHEA 686
 DB 601 EVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNASLNATDKWAFPLHEA 660
 QY 687 AOKGRTQCALLLAHGADPTLKNOEGOTPLDLSADVDVSAALLTAAMPSPALPSYKPKQVL 746
 DB 661 AOKGRTQCALLLAHGADPTLKNOEGOTPLDLSADVDVSAALLTAAMPSPALPSYKPKQVL 720
 QY 747 NGVRSFGATADALSSGSPSSPSLSAASDLNLSGFSSELSVSSVSSGTEGASSLEKKEVP 806
 DB 721 NGVRSFGATADALSSGSPSSPSLSAASDLNLSGFSSELSVSSVSSGTEGASSLEKKEVP 780
 QY 807 GVDFTSQFVRNLGLHMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVER 866
 DB 781 GVDFTSQFVRNLGLHMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVER 840
 QY 867 LISGQOGLNPYLTNTSGSGTILIDLSPDDKEFQSVSEEMQSTVREHROGSHAGGIFNRY 926
 DB 841 LISGQOGLNPYLTNTSGSGTILIDLSPDDKEFQSVSEEMQSTVREHROGSHAGGIFNRY 900
 QY 927 NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNALIHKGFDERHAYIG 986
 DB 901 NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNALIHKGFDERHAYIG 960
 QY 987 GMFGAGIYFAENSCKSNQYVYIGGGTGCPCVHKDRSCYICHROLLFCRVTLGKSFLOFSA 1046
 DB 961 GMFGAGIYFAENSCKSNQYVYIGGGTGCPCVHKDRSCYICHROLLFCRVTLGKSFLOFSA 1020
 QY 1047 MKMAHSPGHHVSTGRPSVNGLAALAEYIYRGEQAYPEYLITYQIMRPEGVMDG 1100
 DB 1021 MKMAHSPGHHVSTGRPSVNGLAALAEYIYRGEQAYPEYLITYQIMRPEGVMDG 1074

RESULT 11
 AAB47022
 ID AAB47022 standard; Protein; 1166 AA.
 XX
 AC AAB47022;
 XX
 DT 29-MAR-2001 (first entry)
 XX
 DE Human SPANK.

XX SPANK: SAM: sterile alpha motif; PARG: insulin resistance;
 KW poly adenosine diphosphate-ribose polymerase; catalytic domain;
 KW ANK: ankyrin repeat; cytosol; insulin-responsive aminopeptidase;
 KW IRAP: GLUT4; adipocyte; insulin signalling pathway; hyperlipidaemia;
 KW glucose intolerance; atheromatous disease; atherosclerosis;
 KW obesity; cardiac insufficiency; coronary insufficiency; stroke;
 KW high blood pressure; non-insulin dependent diabetes; hypertension;
 KW hyperuricaemia; Syndrome X; muscular dystrophy; muscle atrophy.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 173..209
 FT /note= "Encoded by nucleotides 768..878"
 FT Misc-difference 327..362
 FT /note= "Encoded by nucleotides 1230..1337"
 FT Misc-difference 375..398
 FT /note= "Encoded by nucleotides 1374..1445"
 FT Misc-difference 482..524
 FT /note= "Encoded by nucleotides 1695..1823"
 FT Misc-difference 641..677
 FT /note= "Encoded by nucleotides 2172..2282"
 PN WO200077225-A1.
 XX
 PD 21-DEC-2000.
 XX
 PF 09-JUN-2000; 2000WO-US15926.
 XX
 PR 11-JUN-1999; 99US-0138957.
 XX
 PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
 PA (GEOH) GEN HOSPITAL CORP.
 XX
 PI Chi N, Lodish HF;
 XX
 XX WPI: 2001-091404/10.
 DR N-PSDB; AAC85294.
 XX
 PT New insulin signalling protein SPANK, useful for reducing body mass,
 PT glucose intolerance or insulin resistance and for preventing or
 PT treating obesity-related and muscle-related diseases
 XX
 DS Claim 3; Fig 3; 65pp; English.
 XX
 CC This sequence represents human SPANK. The SPANK protein comprises
 CC 3 domains:
 CC (a) a SAM (sterile alpha motif) domain;
 CC (b) a PARG (poly adenosine diphosphate-ribose polymerase) catalytic
 CC domain; and
 CC (c) an ANK domain composed of ankyrin repeats.
 CC SPANK is a cytosolic protein which can poly(ADP-ribose)late itself.
 CC SPANK binds insulin-responsive aminopeptidase (IRAP) and modulates
 CC translocation of GLUT4 in the perinuclear region of adipocytes. It
 CC is an effector in the insulin signalling pathway in eukaryotic cells.
 CC SPANK is useful for reducing body mass, reducing glucose
 CC intolerance or insulin resistance, for preventing or treating
 CC obesity-related diseases or disorders, such as obesity, cardiac
 CC insufficiency, coronary insufficiency, stroke, hypertension,
 CC atheromatous disease, atherosclerosis, high blood pressure, non-insulin
 CC dependent diabetes, hyperlipidaemia, hyperuricaemia and Syndrome X and is
 CC also useful for preventing or treating muscle-related diseases or
 CC disorders, such as muscular dystrophy, muscle atrophy and muscle
 CC fatigue. Antibodies immunospecific for SPANK are useful for detecting
 CC the presence of SPANK polypeptide in a biological sample.
 XX
 SQ Sequence 1166 AA;

Query Match 90.2%; Score 5201; DB 22; Length 1166;
 Best Local Similarity 87.3%; Pred. No. 0;
 Matches 1025; Conservative 1; Mismatches 0; Indels 148; Gaps 10;

QY 1 GFRKDVVEYLQNGASVQARDGGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 60
 DB 67 GFRKDVVEYLQNGASVQARDGGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 126
 QY 61 PLHEAAIKGKIDVICIVLLQHGAEPTIRNTDGTALDLPDSAKAVL----- 106
 DB 127 PLHEAAIKGKIDVICIVLLQHGAEPTIRNTDGTALDLPDSAKAVLKMALLTPLNVNCH 186
 QY 107 ---TCEYKDELLESARSNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVYIQ 162
 DB 187 ASDGTGEYKDELLESARSNEE-----RKSTPLHLAAGYNRVYIQ 228
 QY 163 LLLQHGADVHAKDKGDLVPLHNACSYGHEVTELLVKGACVNMADLMQFTPLHEAASKN 222
 DB 229 LLLQHGADVHAKDKGDLVPLHNACSYGHEVTELLVKGACVNMADLMQFTPLHEAASKN 288
 QY 223 RVVCSLLSYGADPTLLCHNKSASDAIDAPTOLKERL-----AYEFK 266
 DB 289 RVVCSLLSYGADPTLLCHNKSASDAIDAPTOLKERLKKHLSLEWVNFKHPQAYEFK 348
 QY 267 HSLQAAAREADVTRIKKHLSEWVNFKHPQTHETALHCAAAAPYKRRKOICEYLLRKGAN 326
 DB 349 HSLQAAAREADVTR-----THETALHCAAAAS---KRKOICEYLLRKGAN 389
 QY 327 INEKTQ---EFLTPLVHASEKAHNDVVEVVEVVEVVEVVEVVEVVEVVEVVEVVEV 383
 DB 390 INEKTQPYEFLTPLVHASEKAHNDVVEVVEVVEVVEVVEVVEVVEVVEVVEVVEV 449
 QY 384 RLLLSYGCNPNTLSLOGFTALQMGNNVOOLL-----QEGISLGNSE 425
 DB 450 RLLLSYGCNPNTLSLOGFTALQMGNNVOOLLTVKLCVTQSVNCRDIEGQGGISLGNSE 509
 QY 426 ADROLLEAKAGDVETVKKLTQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLLQHGAD 485
 DB 510 ADROLLEAKAGDVE-----ROSTPLHFAAGYNRVSVVEYLLQHGAD 551
 QY 486 VHAOKGGLVPLHNACSYGHEVVAELLVKGAVVAVADLMKFTPLHEAAAKGYEICKLL 545
 DB 552 VHAOKGGLVPLHNACSYGHEVVAELLVKGAVVAVADLMKFTPLHEAAAKGYEICKLL 611
 QY 546 LQHGADPTKKNRGDNTPLDVKDGDTDIQLDLRGDAALLDAAKKGCLARVKKLSPPDYN 605
 DB 612 LQHGADPTKKNRGDNTPLDVKDGDTDIQL-----RVKLSPPDYN 652
 QY 606 CRDTQ-----RHSTPLHLAAGYNNLEVAEYLLQHGADVNAODKGG 646
 DB 653 CRDTQDLRLRGDAALLDAAKKGCLARHSTPLHLAAGYNNLEVAEYLLQHGADVNAODKGG 712
 QY 647 LIPLHNAASGHVDVAALLIKYNACYNATDKWAFPLHEAAQGRTOICALLAHGADPT 706
 DB 713 LIPLHNAASGHVDVAALLIKYNACYNATDKWAFPLHEAAQGRTOICALLAHGADPT 772
 QY 707 LKNQEGQTPDLVSDVADVSALLTAMPSPALSCYKPOVLNGVRSPGATADALSSGSPSP 766
 DB 773 LKNQEGQTPDLVSDVADVSALLTAMPSPALSCYKPOVLNGVRSPGATADALSSGSPSP 832
 QY 767 SLSAASLNLNLSSGFSSELSSVSSSGTEGASLESKEVEKVEPVDFSTQFVNRNLGLEHMD 826
 DB 833 SLSAASLNLNLSSGFSSELSSVSSSGTEGASLESKEVEKVEPVDFSTQFVNRNLGLEHMD 892
 QY 827 IFERQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGOOGLNPYLTLNTSGSG 886
 DB 893 IFERQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGOOGLNPYLTLNTSGSG 952
 QY 887 TILDLSPDDKPEQSVSEEMQSTVREHDDGHAGIENRYNLIKTKVCNKKLWERYTHR 946
 DB 953 TILDLSPDDKPEQSVSEEMQSTVREHDDGHAGIENRYNLIKTKVCNKKLWERYTHR 1012
 QY 947 RKEYSENNHNAERMLFHGSPFVNAIHKGFDERHAYIGMFGAGIYFAENSKSNQYV 1006
 DB 1013 RKEYSENNHNAERMLFHGSPFVNAIHKGFDERHAYIGMFGAGIYFAENSKSNQYV 1072
 QY 1007 YGIGGGTGPVHKDRSCYICHROLLEFCRVTLGKSFLOFSAMKMAHSPPGHSHVTRPSVN 1056

72-1
 seq 10:3
 result 14
 10260/10

```

Db 1073 YGIGGGTGVHDKRSCYICHRQLLCFVTLGKSFLOFSAMKMAHSPPGHRSVTGRPSVN 1132
Qy 1067 GLALAEEVYVIRGEQAYPEYLITYQIMRPGWVDG 1100
Db 1133 GLALAEEVYVIRGEQAYPEYLITYQIMRPGWVDG 1166

RESULT 12
AAB27209
ID AAB27209 standard; Protein; 1333 AA.
XX
AC AAB27209;
XX
DT 27-FEB-2001 (first entry)
XX
DE Human tankyrase II SEQ ID NO: 2.
XX
KW Human; tankyrase II; telomere length; signal transduction.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..180 Domain
FT /note= "ANK domain"
FT 181..1017 Domain
FT /note= "ANK domain"
FT Misc-difference 199
FT /label= Xaa
FT /note= "encoded by CNN"
FT Misc-difference 208..209
FT /label= XaaXaa
FT /note= "encoded by ANCCNC"
FT Misc-difference 212
FT /label= Xaa
FT /note= "encoded by GNG"
FT Misc-difference 232
FT /label= Xaa
FT /note= "encoded by NTT"
FT Misc-difference 242
FT /label= Xaa
FT /note= "encoded by NTC"
FT Misc-difference 244
FT /label= Xaa
FT /note= "encoded by NCA"
FT Misc-difference 255
FT /label= Xaa
FT /note= "encoded by ANT"
FT Misc-difference 264
FT /label= Xaa
FT /note= "encoded by ANT"
FT Misc-difference 726
FT /label= Xaa
FT /note= "encoded by TNT"
FT Misc-difference 794
FT /label= Xaa
FT /note= "encoded by GAN"
FT Misc-difference 800
FT /label= Xaa
FT /note= "encoded by NTG"
FT Misc-difference 822..823
FT /label= XaaXaa
FT 1030..1095 Domain
FT /note= "encoded by TTTNTNT"
FT /note= "SAM domain"
FT Misc-difference 1055
FT /label= Xaa
FT /note= "encoded by NAG"
FT 1190..1328 Domain
FT /note= "Xaa-unknown"
FT /note= "PARP domain"
XX
PN WO20061813-A1.

```

```

XX 19-OCT-2000.
XX
XX 10-APR-2000; 2000WO-US09558.
XX
XX 09-APR-1999; 99US-0128577.
XX 13-APR-1999; 99US-0129123.
XX (GERO-) GERON CORP.
XX
XX PA Morin GB, Funk WD, Piatyszek MA;
XX PI WPI; 2000-679503/66.
XX DR N-PSDB; AAC66823.
XX
XX Novel mammalian Tankyrase II polypeptide and the polynucleotide
XX encoding the polypeptide useful for modulating or maintaining telomere
XX length, replicative capacity, apoptosis, chromosome packing or gene
XX expression
XX
XX Disclosure; Fig 2; 52pp; English.
XX
XX The present sequence is a version of the human tankyrase II protein
XX sequence. The protein is thought to be involved in signal transduction in
XX the cell, and to have binding activity for other telomere-associated
XX proteins. It is possible that it plays a role in the regulation of
XX telomere length, thus affecting the replicative ability of the cell. The
XX protein is useful for ribosylating target proteins, for determining
XX tankyrase II binding activity in a sample, and for modulating telomere
XX length in a cell.
XX
XX Sequence 1333 AA;
XX
Query Match 85.5%; Score 4932.5; DB 21; Length 1333;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 963; Conservative 8; Mismatches 129; Indels 9; Gaps 2;
QY 1 GFGRKDVVYLLQNGASVQARDGGLIPLHNACSFHGAEVYVLLRHGADPNARDNWNVT 60
DB 225 GFGRKLNKXLYLLQNGANXQLYNGGLIPLHXACSFHGAHXKINLLRHGADPNARDNWNVT 284
QY 61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKKDELLESAR 120
DB 285 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKKDELLESAR 344
QY 121 SGNEEKWMLTPLNVNCHASDGRKSTPLHLAAGYNVYKIVOLLQHGADYHAKDKGLV 180
DB 345 SGNEEKWMLTPLNVNCHASDGRKSTPLHLAAGYNVYKIVOLLQHGADYHAKDKGLV 404
QY 181 PLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVVCSSLISYGADPTLL 240
DB 405 PLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVVCSSLISYGADPTLL 464
QY 241 NCHNKSALDAPTOLKERLAYEFKHSLLQAAREADVTRIKKHLSEMVNFKHPQTHET 300
DB 465 NCHNKSALDAPTOLKERLAYEFKHSLLQAAREADVTRIKKHLSEMVNFKHPQTHET 524
QY 301 ALHCAAAAPYKRRKQICELLRRKANINEKTEPLTHVASEKAHNDVYVVVVKHEAV 360
DB 525 ALHCAAAAPYKRRKQICELLRRKANINEKTEPLTHVASEKAHNDVYVVVVKHEAV 584
QY 361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNILISLOGFTALQMGNNVQQLLEGIS 420
DB 585 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNILISLOGFTALQMGNNVQQLLEGIS 644
QY 421 LGNSEADROLLEAAKAGDVETVKKCTVQSVNCHDIEGROSTPLHFAAGNRYSVVEYLL 480
DB 645 LGNSEADROLLEAAKAGDVETVKKCTVQSVNCHDIEGROSTPLHFAAGNRYSVVEYLL 704
QY 481 QHGADYHAKDKGLVPLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRV 540
DB 705 QHGADYHAKDKGLVPLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRV 764

```

Qy 541 ICKLLQHGADPTKKNRDNTPDLVDGDTDIODLLRGDAALLDAAKGCLARVKKLSS 600
 Db 765 ICKLLQHGADPTKKNRDNTPDLVDGDTDIODLLRGDAALLDAAKGCLARVKKLSS 824
 Qy 601 PDVNCRDTOGRHSTPLHLA--AGYNNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGH 658
 Db 825 PDVNCRDTOGRHSTPLHLAAGXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 884
 Qy 659 VDVAALLIKYNACVNATDKWAFPLHEAAQGRQLCALLLAHGADPTLKNQEQOTPLDL 718
 Db 885 XXX 944
 Qy 719 VSADVSALLTAAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASSLDNL 778
 Db 945 XXXXXXXXXXXTAAMPSPVLPSCNKPVLNGVRSFGATADALSSGSPSSLSAASSLDNL 1004
 Qy 779 SGFSFELSSVSSSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHMDIFEREQITLDVL 838
 Db 1005 SGFSFELSSVSSSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHMDIFEREQITLDVL 1064
 Qy 839 VEMGHKELKEIGINAYGHRHKLIGVERLISGOQGLNPYLTNTSGSGTILIDLSPODKE 898
 Db 1065 VEMGHKELKEIGINAYGHRHKLIGVERLISGOQGLNPYLTNTSGSGTILIDLSPODKE 1124
 Qy 899 FQSVSEEMQSTVREHRDGHAGGIFNRYNLIKIOKVCNK-----KLWERYTHRRKEVS 951
 Db 1125 FQSVSEEMQSTVREHRDGHAGGIFNRYNLIKIOKVCNKXXXXXAKIRHEERYTHRRKEVS 1184
 Qy 952 EENNHNANERMLFHGSPFVNATIIHKGFDERHAYTGGMFGAGIYFAENSCKSNQVYVGTG 1011
 Db 1185 EENNHNANERMLFHGSPFVNATIIHKGFDERHAYTGGMFGAGIYFAENSCKSNQVYVGTG 1244
 Qy 1012 GTGCPVHKDRSCYICHROLLFCRVTLGKSFLOFSAMKMAHSPGPHSHSVTGPRSVNGLALA 1071
 Db 1245 GTGCPVHKDRSCYICHROLLFCRVTLGKSFLOFSAMKMAHSPGPHSHSVTGPRSVNGLALA 1304
 Qy 1072 EVVIYRGEQAYPEYLITYQIMRPEGMDVG 1100
 Db 1305 EVVIYRGEQAYPEYLITYQIMRPEGMDVG 1333
 RESULT 13
 AAB27212
 ID AAB27212 standard; Protein; 1327 AA.
 AC AAB27212;
 AC AAB27212;
 DT 27-FEB-2001 (first entry)
 DE Human tankyrase I protein sequence SEQ ID NO: 8.
 KW Human; tankyrase II; telomere length; signal transduction.
 OS Homo sapiens.
 PN WO200061813-A1.
 PD 19-OCT-2000.
 PF 10-APR-2000; 2000WO-US09558.
 PR 09-APR-1999; 99US-0128577.
 PR 13-APR-1999; 99US-0129123.
 PA (GERO-) GERON CORP.
 PI Morin GB, Funk WD, Piatyszek MA;
 DR WPI; 2000-679503/66.
 PT Novel mammalian Tankyrase II polypeptide and the polynucleotide
 PT encoding the polypeptide useful for modulating or maintaining telomere
 PT length, replicative capacity, apoptosis, chromosome packing or gene

expression -
 Claim 4; Fig 5; 52pp; English.
 The present sequence is a version of the human tankyrase I protein
 sequence. The invention relates to the isolation of the protein and
 coding sequences of human tankyrase II. This protein is thought to be
 involved in signal transduction in the cell, and to have binding activity
 for other telomere-associated proteins. It is possible that it plays a
 role in the regulation of telomere length, thus affecting the replicative
 ability of the cell. The protein is useful for ribosylating target
 proteins, for determining tankyrase II binding activity in a sample, and
 for modulating telomere length in a cell.
 SQ Sequence 1327 AA;
 Query Match 84.8%; Score 4889.5; DB 21; Length 1327;
 Best Local Similarity 83.0%; Pred. No. 0;
 Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;
 Qy 1 GFRKDVVEYLLQNGASVOARDGGGLPLHNACSFHAEVNVNLLRHGADPNARDNNWT 60
 Db 225 GFRKDVVEYLLQNGASVOARDGGGLPLHNACSFHAEVNVNLLRHGADPNARDNNWT 284
 Qy 61 PLHEAATKGIKIDYCVILVQHGAEPTIRNTDGTALDADPSAKAVLTGEYKDELLESAR 120
 Db 285 PLHEAATKGIKIDYCVILVQHGAEPTIRNTDGTALDADPSAKAVLTGEYKDELLESAR 344
 Qy 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRKIVOLLQHGADVHAKDGDLV 180
 Db 345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRKIVOLLQHGADVHAKDGDLV 404
 Qy 181 PLHNACSYGHEVTELLVKGACVNAWDLQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
 Db 405 PLHNACSYGHEVTELLVKGACVNAWDLQFTPLHEAASKNRVEVCSLLLSYGADPTLV 464
 Qy 241 NCHNKAIDLAPTOLKERLAYEFKGHSLLOAAAREADVTRIKKHLSELMVNFKHPOHET 300
 Db 465 NCHGKSADVNAPTPELNERUTYEFKGHSLLOAAAREADVTRIKKHLSELMVNFKHPOHET 524
 Qy 301 ALHCAASAPYKPKQICELLRLKCANINEKTKEFTPLHVASEKAHNDVVEVVKHEAV 360
 Db 525 ALHCAVASLHPKPKQVTELLRLKCANINEKTKEFTPLHVASEKAHNDVVEVVKHEAV 584
 Qy 361 NALDNLGOTSILHRAAYCGHLQTCRLLSYSGCDPNIISLOGFTALQMGNNENVQILSESTP 420
 Db 585 NALDNLGOTSILHRAAYCGHLQTCRLLSYSGCDPNIISLOGFTALQMGNNENVQILSESTP 644
 Qy 421 LGNSEADQLLEAAKAGDVEVTKKLTQVSNCRDIEGROSTPLHFAAGYNRVSVVEYLL 480
 Db 645 IRTSDVDYRLLEASKAGDLETQKLCSSQNVNCRDLRGHSTPLHFAAGYNRVSVVEYLL 704
 Qy 481 QHGADVHAKDGGLVPLHNACSYGHEVTELLVKGACVNAWDLQFTPLHEAASKNYE 540
 Db 705 HGGADVHAKDGGLVPLHNACSYGHEVTELLVKGACVNAWDLQFTPLHEAASKNYE 764
 Qy 541 ICKLLQHGADPTKKNRDNTPDLVDGDTDIODLLRGDAALLDAAKGCLARVKKLSS 600
 Db 765 ICKLLQHGADPTKKNRDNTPDLVDGDTDIODLLRGDAALLDAAKGCLARVKKLSS 824
 Qy 601 PDVNCRDTOGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGH 660
 Db 825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGH 884
 Qy 661 VAALLIKYNACVNATDKWAFPLHEAAQGRQLCALLLAHGADPTLKNQEQOTPLDLS 720
 Db 885 IAALLIKYNACVNATDKWAFPLHEAAQGRQLCALLLAHGADPTLKNQEQOTPLDLS 944
 Qy 721 ADDVSALLTAAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASSLDNL 780
 Db 945 ADDIRALLIDAMPPEALPTCFKPOAT-----VVSASLISPASTPSCLSAASSLDNL 997
 Qy 781 SFSELSSVSSSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHMDIFEREQITLDVL 838

Db 998 PLAEAVGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIPEFETQITLDVL 1057
QY 839 VEMGHKEKEIGINAYGHRHKLKGVRLISGOGLNPYLTLNTSGSGTILIDLPDDKE 898
Db 1058 ADMGHEELKEIGINAYGHRHKLKGVRLISGOGLNPYLTLNTSGSGTILIDLPDDKE 1117
QY 899 FOSVEEMQSTVREHRRDGGHAGGIFNRYNLIKIQKVCNKKLWERYTHRRKEVSEENHHA 958
Db 1118 YQSVEEMQSTVREHRRDGGHAGGIFNRYNLIKIQKVCNKKLWERYTHRRKEVSEENHHA 1177
QY 959 NERMLFHGSPFVNAIIHKGFDERHAYTGGMFGAGIYFAENSSKSNQYVYGGTGCPVH 1018
Db 1178 NERMLFHGSPFVNAIIHKGFDERHAYTGGMFGAGIYFAENSSKSNQYVYGGTGCPVH 1237
QY 1019 KDRSCYICHRQLLCFVRTYLGKSLFQFSAMKMAHSPGCHSHVTCRPSVNGLALEYVYRG 1078
Db 1238 KDRSCYICHRQLLCFVRTYLGKSLFQFSAMKMAHSPGCHSHVTCRPSVNGLALEYVYRG 1297
QY 1079 EOAYPEYLITYQIMRPE 1095
Db 1298 EOAYPEYLITYQIMKPE 1314

RESULT 14

AA144402
ID AAY44402 standard; protein; 1327 AA.

AC AAY44402;

DT 22-MAR-2000 (first entry)

XX Human tankyrase.

XX Human tankyrase; TRF1, telomeric repeat binding factor-1;
KW two-hybrid screen; telomere length regulation; recombinant tankyrase;
KW aging; skin atrophy; macular degeneration; atherosclerosis;
KW ataxia telangiectasia; tankyrase modulator; rational drug design.
OS Homo sapiens.

XX Key Location/Qualifiers
FH 181..1010
FT Domain /note= "Ankyrin-specific (ANK) repeat consensus domain"
FT 1023..1088
FT /label= Sterile_alpha_motif
FT 1176..1314
FT /note= "Poly(ADP-ribose)polymerase (PARP) related domain"

XX WO9964606-A1.

XX 16-DEC-1999.

XX 09-JUN-1999; 99WO-US12968.

XX 10-JUN-1998; 98US-0095225.

XX 17-AUG-1998; 98US-0135233.

XX 19-NOV-1998; 98US-0196387.

XX (UYRQ) UNIV ROCKEFELLER.

XX De Lange T, Smith S;

XX WPI; 2000-116549/10.

XX N-PSDB; AAZ29627.

XX New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT length, used to identify modulators, e.g. for inhibiting growth of
PT cancer
XX Claim 1; Fig 1B; 133pp; English.
XX The present sequence is human tankyrase, a TRF1 (telomeric repeat binding

CC factor-1) binding protein. This was isolated using a two-hybrid screen
CC with TRF1. Tankyrase is involved in regulation of telomere length through
CC TRF1 and may directly modulate the effect of TRF1. Recombinant tankyrase,
CC or its fragments are used to identify specific modulators which are
CC potential drugs for counteracting telomere shortening associated with aging
CC (e.g. atrophy of skin, macular degeneration or atherosclerosis) or
CC diseases like ataxia telangiectasia. They are also used to raise specific
CC antibodies (used as immunoassay reagents and as modulators of tankyrase
CC activity) and in rational drug design.

XX Sequence 1327 AA;

Query Match 84.8%; Score 4889.5; DB 21; Length 1327;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

QY 1 GFRKKDVEYLLONGASVOARDGGLIPLHNACSFHAEVYVNLRLRHGADPNARDNNYIT 60
Db 225 GFRKKDVEYLLONGASVOARDGGLIPLHNACSFHAEVYVNLRLRHGADPNARDNNYIT 284
QY 61 PLHEAAIKGKIDVICIVLLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKKDELLESAR 120
Db 285 PLHEAAIKGKIDVICIVLLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKKDELLESAR 344
QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLV 180
Db 345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLV 404
QY 181 PLHNACSYGHYEYVELLVKHGACVNMADLWQFTPLHEAASKNRVEYCSLLSHGADPTLL 240
Db 405 PLHNACSYGHYEYVELLVKHGACVNMADLWQFTPLHEAASKNRVEYCSLLSHGADPTLL 464
QY 241 NCHNKSATDLAPTOLKERYLAEFEKSHLSLOAREADVTIRIKHLSLEMYNFKHPQTHET 300
Db 465 NCHNKSATDLAPTOLKERYLAEFEKSHLSLOAREADVTIRIKHLSLEMYNFKHPQTHET 524
QY 301 ALHCAASPPYKPKQICELLRRKANINEKTEFLTPHVASEKAHNDVYVYVVKHEAKV 360
Db 525 ALHCAASPPYKPKQICELLRRKANINEKTEFLTPHVASEKAHNDVYVYVVKHEAKV 584
QY 361 NALDNLGQISLHRAAYCGHLOTCTRLLSYCGDNPITISLQGTALQMGNEVVOQLQSGIS 420
Db 585 NALDNLGQISLHRAAYCGHLOTCTRLLSYCGDNPITISLQGTALQMGNEVVOQLQSGIS 644
QY 421 LGSEADROLLEAAKAGDVETVKKLCITQSVNCRDLLEGROSTPLHFAAGYVYVVEYLL 480
Db 645 IRTSDVDYRLLEAKAGDVETVKKLCITQSVNCRDLLEGROSTPLHFAAGYVYVVEYLL 704
QY 481 QHGADVHAKDKGLVPLHNACSYGHYEYVELLVKHGACVNMADLWQFTPLHEAASKNRVEY 540
Db 705 QHGADVHAKDKGLVPLHNACSYGHYEYVELLVKHGACVNMADLWQFTPLHEAASKNRVEY 764
QY 541 ICKLLLOHGADPTKKNRDGNTPLDLVKDGTDTQDILLRGDAALLDAAKGCLARVKKLS 600
Db 765 ICKLLLOHGADPTKKNRDGNTPLDLVKDGTDTQDILLRGDAALLDAAKGCLARVKKLS 824
QY 601 PDVNCNCRDTQGRHSTPLHLAAGYVYVVEYLLQHGADVNAQDKGLIPLHNAASYGHVD 660
Db 825 PDVNCNCRDTQGRHSTPLHLAAGYVYVVEYLLQHGADVNAQDKGLIPLHNAASYGHVD 884
QY 661 VAALLIKYNACVNATDKWAFTPLHEAASKNRVEYVELLVKHGACVNMADLWQFTPLHEAASKNRVEY 720
Db 885 VAALLIKYNACVNATDKWAFTPLHEAASKNRVEYVELLVKHGACVNMADLWQFTPLHEAASKNRVEY 944
QY 721 ADDVSALLTAAMPSPSALPCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLNLG 780
Db 945 ADDVSALLTAAMPSPSALPCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLNLG 997
QY 781 SFSEISSVSSSGTEGASSLEKK--EVPGVDFSTITQVFNGLGHEHMDIFEREQITLDVL 838
Db 998 PLAEAVGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIPEFETQITLDVL 1057
QY 839 VEMGHKEKEIGINAYGHRHKLKGVRLISGOGLNPYLTLNTSGSGTILIDLPDDKE 898

```

Db 1058 ADMGHEELKEIGNAYGHRHKLKIGVERLLGGQGTNPYLTFCVNOGTILLDLAPEDKE 1117
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 899 FOSVEEEMQSTVREHROGGHAGGIFNRYNLTQKVCNKKLWERYTHRRKEVSEENHNHA 958
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1118 YQSVVEEEMQSTIREHROGGNAGGIFNRYNIRIQKVNNKKLRFERFCHROKQEVSEENHNH 1177
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 959 NERMLFHGSPFVNAIHKGFDERHAYIGGMPGAGIYFAENSCKSNQYVYGIGGTGCPVH 1018
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1178 NERMLFHGSPFNAIHKGFDERHAYIGGMPGAGIYFAENSCKSNQYVYGIGGTGCPH 1237
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1019 KDRSCYCHROLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGIALAEYIYRG 1078
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1238 KDRSCYCHROLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGIALAEYIYRG 1297
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1079 EOAYPEYLITYOIMRPE 1095
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1298 EOAYPEYLITYOIMKPE 1314
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

RESULT 15

AAB66279 standard; Protein: 1327 AA.

XX ID AAB66279;

XX AC AAB66279;

XX DT 05-APR-2001 (first entry)

XX DE Human tankyrase1 SEQ ID NO: 4.

XX KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

XX KW inflammatory disorder.

XX OS Homo sapiens.

XX PN WO200100849-A1.

XX PD 04-JAN-2001.

XX PF 28-JUN-2000; 2000WO-US17827.

XX PR 29-JUN-1999; 99US-0141582.

XX PA (ICOS-) ICOS CORP.

XX PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX DR WPI; 2001-102896/11.

XX DR N-PSDB; AAF63838.

XX PT New tankyrase2 polypeptides, useful for treating conditions mediated by
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
 PT inflammatory and autoimmune disorders -

XX PS Example 1; Page 118-121; 242pp; English.

XX CC The present invention provides the protein and coding sequence for the
 CC human tankyrase2 protein. This is found in two different versions,
 CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
 CC polyADP-ribosylation activity and is involved in the modification of
 CC TRF1, which is a telomere-specific binding protein. The regulation of
 CC telomere length, in which TRF1 has a role, is linked to ageing and
 CC cancer. The sequences are useful in the treatment of cancers and
 CC inflammatory disorders.

XX SQ Sequence 1327 AA;

Query Match 84.8%; Score 4889.5; DB 22; Length 1327;

Best Local Similarity 83.0%; Pred. No. 0;

Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

Qy 1 GFGKRDVVEYLLONGASVOARDGGLIPLHNACSFCHAEVFNLLRHGADPNARDNWNVT 60
 :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

Db 225 GFGKRDVVEYLLONGASVOARDGGLIPLHNACSFCHAEVFNLLRHGADPNARDNWNVT 284
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 61 PLHEAAIKGKIDVCIVLQHGAEPTIRNTDGTALDLDPSAKAVLTGYKKDELLESAR 120
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 285 PLHEAAIKGKIDVCIVLQHGADPNIRNTDGRSALDLDPSAKAVLTGYKKDELLESAR 344
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGDLV 180
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGGLV 404
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 181 PLHNACSYGHEVTELLVYKHGACVNAWLQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 405 PLHNACSYGHEVTELLVYKHGACVNAWLQFTPLHEAASKNRVEVCSLLLSYGADPTLV 464
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 241 NCHNKSAIDLPTPOLKERLAYEFKGHSLQAARADVTIRIKKHLSELMVNEKHPQTHET 300
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 465 NCHGKSAYDMAPTPELRERLYEFKGHSLQAARADLAKVKT LALEIINFKPQSHET 524
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 301 ALHCAASPYPRKQICELLRGANINEKTEFLPLHVASEKAHNVDVVEVVKHEAKV 360
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 525 ALHCAVASLHPRKQVTELLRKGANVNEKNDFTPLHVAAERAHNDVNEVLHKGAKM 584
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 361 NALDNLGOTSLHRAAYCGHLQTCRLLLSYCGDPNIISLQFTALOMGNENVOOLLQEGIS 420
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 585 NALDTLGTALHRAALAGHLQTCRLLLSYCGDPSIISLQFTAAOMGNENVOOIIUSESTP 644
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 421 LGNSEADROLLEAAKAGDVETVKKLCVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDIEGRHSTPLHFAAGYNRVSVVEYLL 704
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 481 QHGADVHAKDKGGLVPLHNACSYGHEVYAEVLVYKHGADVNAQDKGGLIPLHNAASYGHVD 540
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 705 HHGADVHAKDKGGLVPLHNACSYGHEVYAEVLVYKHGADVNAQDKGGLIPLHNAASYGHVD 764
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 541 ICKLLQHGADPTKKNRDGNTPLDLVKDGTDIQDLRLGGDAALLDAKKGCLARVKLSS 600
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 765 ICKLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLRLGGDAALLDAKKGCLARVKLCT 824
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 601 PDNVNCRDITQGRHSTPLHLAAGYNVLEAEVYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 825 PENINCEDTQGRNSTPLHLAAGYNVLEAEVYLLQHGADVNAQDKGGLIPLHNAASYGHVD 884
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 661 VAALLTKYNACVNATDKWAFTPLHEAAQKQRTQLCALLLAHGADPTLKNOEGQTPDLVLS 720
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 885 IAALLIKYNTCVNATDKWAFTPLHEAAQKQRTQLCALLLAHGADPTLKNOEGQTPDLAT 944
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 721 ADVSALLTAAPPSPALPCYKQVNLGVSPGATADALSSGSPSPSSLSAASLDNLG 780
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 945 ADDIRALLIDAMPPEALPTCFKQAT-----VVSASLISPASTFSCLSAASIDNLG 997
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 781 SFSELSSWSSSGTSGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDFEREQITLDVL 838
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 998 PLAEVAVGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDFETBOITLDVL 1057
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 839 VEMGHKEIKEIGNAYGHRHKLKIGVERLLGGQGTNPYLTFCVNOGTILLDLAPEDKE 898
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1058 ADMGHEELKEIGNAYGHRHKLKIGVERLLGGQGTNPYLTFCVNOGTILLDLAPEDKE 1117
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 899 FOSVEEEMQSTVREHROGGHAGGIFNRYNLTQKVCNKKLWERYTHRRKEVSEENHNHA 958
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1118 YQSVVEEEMQSTIREHROGGNAGGIFNRYNIRIQKVNNKKLRFERFCHROKQEVSEENHNH 1177
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 959 NERMLFHGSPFVNAIHKGFDERHAYIGGMPGAGIYFAENSCKSNQYVYGIGGTGCPVH 1018
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1178 NERMLFHGSPFNAIHKGFDERHAYIGGMPGAGIYFAENSCKSNQYVYGIGGTGCPH 1237
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1019 KDRSCYCHROLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGIALAEYIYRG 1078
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1238 KDRSCYCHROLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGIALAEYIYRG 1297
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1079 EOAYPEYLITYOIMRPE 1095
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1298 EOAYPEYLITYOIMKPE 1314
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```


Wed Feb 12 16:08:51 2003

us-09-843-159b-3.rag

Page 16

Search completed: February 12, 2003, 03:32:18
Job time : 43.6068 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 07:42:50 : Search time 15,0427 Seconds
(without alignments)
2151.551 Million cell updates/sec

Title: US-09-843-159B-3

Perfect score: 5769

Sequence: 1 GFGKRDVVEYLQNGASVQA.....AYPEYLITYQIMRPEGMVDG 1100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pap.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pap.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pap.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pap.*
- 5: /cgn2_6/ptodata/2/iaa/PCFUS_COMB.pap.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5717	99.1	1166	4	US-09-350-982C-5
2	4889.5	84.8	1327	4	US-09-196-387-2
3	3093	53.6	949	4	US-09-196-387-10
4	1918.5	33.3	673	4	US-09-196-387-8
5	860.5	14.9	1839	2	US-09-172-977-4
6	841.5	14.6	843	2	US-09-172-977-3
7	821	14.2	1745	2	US-09-031-485-33
8	821	14.2	1745	2	US-08-847-429A-33
9	821	14.2	1745	3	US-09-065-474-33
10	821	14.2	1745	4	US-09-557-034-33
11	629.5	10.9	1088	4	US-09-082-059-2
12	513	8.9	352	3	US-09-065-474-139
13	513	8.9	352	4	US-09-557-034-139
14	421.5	7.3	303	2	US-09-031-485-23
15	421.5	7.3	303	2	US-08-847-429A-23
16	421.5	7.3	303	3	US-09-065-474-23
17	421.5	7.3	303	4	US-09-557-034-23
18	409	7.1	302	2	US-09-031-485-38
19	409	7.1	302	2	US-08-847-429A-38
20	409	7.1	302	3	US-09-065-474-38
21	409	7.1	302	4	US-09-557-034-38
22	392	6.8	348	2	US-09-031-485-28
23	392	6.8	348	2	US-08-847-429A-28
24	392	6.8	348	3	US-09-065-474-28
25	392	6.8	348	4	US-09-557-034-28
26	390.5	6.8	787	4	US-09-188-930-334
27	379.5	6.6	1719	4	US-09-605-785-378

ALIGNMENTS

RESULT 1

US-09-350-982C-5

: Sequence 5, Application US/09350982C

: Patent No. 6455290

: GENERAL INFORMATION:

: APPLICANT: Bertheisen, Jens

: APPLICANT: Toma, Salvatore

: APPLICANT: Isacchi, Antonella

: TITLE OF INVENTION: Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods R

: TITLE OF INVENTION: Same

: FILE REFERENCE: PHRM-0043

: CURRENT APPLICATION NUMBER: US/09/350,982C

: CURRENT FILING DATE: 1999-07-09

: NUMBER OF SEQ ID NOS: 10

: SOFTWARE: PatentIn version 3.0

: SEQ ID NO 5

: LENGTH: 1166

: TYPE: PRT

: ORGANISM: Artificial

: NAME/KEY: misc_feature

: OTHER INFORMATION: Xaa is any amino acid

: NAME/KEY: misc_feature

: LOCATION: (1102)..(1102)

: OTHER INFORMATION: n is any nucleic acid

: NAME/KEY: misc_feature

: LOCATION: (2650)..(2650)

: OTHER INFORMATION: n is any nucleic acid

: US-09-350-982C-5

Query Match 99.1%; Score 5717; DB 4; Length 1166;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1091; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy	1	GFGKRDVVEYLQNGASVQA	RDGGLPLHNACSF	GAEVNVNLLRHGAD	NARDNNYT	60
Db	67	GFGKRDVVEYLQNGASVQA	RDGGLPLHNACSF	GAEVNVNLLRHGAD	NARDNNYT	126
Qy	61	PLHEAAIKGKIDVCI	VLQHGAEPTIRNTD	GTALDLP	SAKAVLTGEYKKDELLESAR	120
Db	127	PLHEAAIKGKIDVCI	VLQHGAEPTIRNTD	GTALDLP	SAKAVLTGEYKKDELLESAR	186
Qy	121	SGNEEKMMALLT	PLNVNCHASDGRKST	PLHLAAGYNRVK	IVQLLQHGADVHAKDKGLV	180
Db	187	SGNEEKMMALLT	PLNVNCHASDGRKST	PLHLAAGYNRVK	IVQLLQHGADVHAKDKGLV	246
Qy	181	PLHNACSYGHVEVTELL	VKKGACVNM	DLWQFTPL	HEAASKNRVEVCSLLSYGADPTLL	240
Db	247	PLHNACSYGHVEVTELL	VKKGACVNM	DLWQFTPL	HEAASKNRVEVCSLLSYGADPTLL	306

QY 241 NCHNKSALDAPTPOLKERLAYEFKHSLLQAAREADVTRIKKHLSELMVNFKHPOTHET 300
Db 307 NCHNKSALDAPTPOLKERLAYEFKHSLLQAAREADVTRIKKHLSELMVNFKHPOTHET 366
QY 301 ALHCAASPPKPKQICELLRRKGANINTEKFLTPHVAASEKAHNDVVEVVKHEAKV 360
Db 367 AXHCAASPPKPKQICELLRRKGANINTEKFLTPHVAASEKAHNDVVEVVKHEAKV 426
QY 361 NALDNLGOTSLSHRAAYCGHLOTCLRLLSYGCDPNITISLOGFTALOMGNENVOQLQEGIS 420
Db 427 NALDNLGOTSLSHRAAYCGHLOTCLRLLSYGCDPNITISLOGFTALOMGNENVOQLQEGIS 486
QY 421 LGNSEADROLLEAAKAGDVETVKKLTCTQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 480
Db 487 LGNSEADROLLEAAKAGDVETVKKLTCTQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 546
QY 481 OHGADVHAKDKGLVPLHNAACSYGHEVABLLVKGAVNVADLWFTPLHFAAAGKYE 540
Db 547 OHGADVHAKDKGLVPLHNAACSYGHEVABLLVKGAVNVADLWFTPLHFAAAGKYE 606
QY 541 ICKLLQHGADPTKKNRDGNTPLDLVKDGTDTIDLLRGDAALLDAKGLARVKKLSS 600
Db 607 ICKLLQHGADPTKKNRDGNTPLDLVKDGTDTIDLLRGDAALLDAKGLARVKKLSS 666
QY 601 PDNVNCRDTQGRHSTPLHIAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYCHVD 660
Db 667 PDNVNCRDTQGRHSTPLHIAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYCHVD 726
QY 661 VAALLIKYNACVNATDKWFTPLHFAAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYCHVD 720
Db 727 VAALLIKYNACVNATDKWFTPLHFAAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYCHVD 786
QY 721 ADVSALLTAAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASSLDNLG 780
Db 787 ADVSALLTAAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASSLDNLG 846
QY 781 SFSELSVSSVSSGTEGASLEKEVPDGFSTQFVNRNLGLEHMDIFEREQITLDVLE 840
Db 847 SFSELSVSSVSSGTEGASLEKEVPDGFSTQFVNRNLGLEHMDIFEREQITLDVLE 906
QY 841 MGHKELKEIGINAYGHRHKLKIGVERLSIQOGLNPLTLNTSGSTIILDSPDKFEQ 900
Db 907 MGHKELKEIGINAYGHRHKLKIGVERLSIQOGLNPLTLNTSGSTIILDSPDKFEQ 966
QY 901 SVEEEMQSTVREHRDGGHAGGIFNRYNLIKIQVCNKKLWERYTHRRKEVSEENHNHANE 960
Db 967 SVEEEMQSTVREHRDGGHAGGIFNRYNLIKIQVCNKKLWERYTHRRKEVSEENHNHANE 1026
QY 961 RMLFHGSPFNALIHKGFDERHAYIGGMFGAGIYFAENSCKSNQYVYIGGTCGCPVHKD 1020
Db 1027 RMLFHGSPFNALIHKGFDERHAYIGGMFGAGIYFAENSCKSNQYVYIGGTCGCPVHKD 1086
QY 1021 RSCYICHRQLLCFRTVLGKSFLOFSAMKMAHSPGHHSHVTRPSVNGLALAEVVIYRGQ 1080
Db 1087 RSCYICHRQLLCFRTVLGKSFLOFSAMKMAHSPGHHSHVTRPSVNGLALAEVVIYRGQ 1146
QY 1081 AYPEYLITYQIMRPGMVDG 1100
Db 1147 AYPEYLITYQIMRPGMVDG 1166

RESULT 2

US-09-196-387-2

; Sequence 2, Application US/09196387

; Patent No. 6277613

; GENERAL INFORMATION:

; APPLICANT: de Lange, Titia

; APPLICANT: Smith, Susan

; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS

; TITLE OF INVENTION: OF USE THEREOF

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/196,387
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/095,225
FILING DATE: June 10, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1327 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-196-387-2

Query Match

Best Local Similarity 83.0%; Score 4889.5; DB 4; Length 1327;

Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

QY 1 GFRKDVVEYLLQNGASVQARDGGIPLHNAACSYGHEVABLLVKGAVNVADLWFTPLHFAAAGKYE 60
Db 225 GFRKDVVEYLLQNGASVQARDGGIPLHNAACSYGHEVABLLVKGAVNVADLWFTPLHFAAAGKYE 284
QY 61 PLHEAAIKGIDVCIVLLOHGAETIRNTDGTALDAPSAKAVLTGEYKDELLESAR 120
Db 285 PLHEAAIKGIDVCIVLLOHGAETIRNTDGTALDAPSAKAVLTGEYKDELLESAR 344
QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHAAAGYNRVRIQVOLLQHGADVHAKDKGLV 180
Db 345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHAAAGYNRVRIQVOLLQHGADVHAKDKGLV 404
QY 181 PLHNAACSYGHEVTELLVKGACVNAQDKGGLIPLHNAACSYGHEVABLLVKGAVNVADLWFTPLHFAAAGKYE 240
Db 405 PLHNAACSYGHEVTELLVKGACVNAQDKGGLIPLHNAACSYGHEVABLLVKGAVNVADLWFTPLHFAAAGKYE 464
QY 241 NCHNKSALDAPTPOLKERLAYEFKHSLLQAAREADVTRIKKHLSELMVNFKHPOTHET 300
Db 465 NCHNKSALDAPTPOLKERLAYEFKHSLLQAAREADVTRIKKHLSELMVNFKHPOTHET 524
QY 301 ALHCAASPPKPKQICELLRRKGANINTEKFLTPHVAASEKAHNDVVEVVKHEAKV 360
Db 525 ALHCAASPPKPKQICELLRRKGANINTEKFLTPHVAASEKAHNDVVEVVKHEAKV 584
QY 361 NALDNLGOTSLSHRAAYCGHLOTCLRLLSYGCDPNITISLOGFTALOMGNENVOQLQEGIS 420
Db 585 NALDNLGOTSLSHRAAYCGHLOTCLRLLSYGCDPNITISLOGFTALOMGNENVOQLQEGIS 644
QY 421 LGNSEADROLLEAAKAGDVETVKKLTCTQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 480
Db 645 LGNSEADROLLEAAKAGDVETVKKLTCTQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 704
QY 481 OHGADVHAKDKGLVPLHNAACSYGHEVABLLVKGAVNVADLWFTPLHFAAAGKYE 540
Db 704 OHGADVHAKDKGLVPLHNAACSYGHEVABLLVKGAVNVADLWFTPLHFAAAGKYE 764

Db 705 HHGADVHAKDGGGLVPLHNACSYGHEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764
QY 541 ICKLLQHGADPTKKNRGNTPLDLVKDGDTDIDQLLRGDAALLDAKKGCLARVKKLSS 600
Db 765 ICKLLKHGADPTKKNRGNTPLDLVKDGDTDIDQLLRGDAALLDAKKGCLARVKKLCT 824
QY 601 PDVNCRTDQGRHSTPLHLAAGYNLEVAEYLLQHGADVNADKGGGLPLHNAASYGHVD 660
Db 825 PENINCRTDQGRNSTPLHLAAGYNLEVAEYLLHGHGADVNADKGGGLPLHNAASYGHVD 884
QY 661 VAALLIKYNACVNTADKWAFTPLHEAAQKGRTOCALLLAHGADPTLKNQEGQTPDLVLS 720
Db 885 IAALLIKYNTCVNADKWAFTPLHEAAQKGRTOCALLLAHGADPTLKNQEGQTPDLVLS 944
QY 721 ADDVALLTAAMPSPSALSCYKQVPLNGVRSPCATADALSSGSPSSLSAASSLDNLGS 780
Db 945 ADDIRALLIDAMPPEALPTCFKQAT-----VVSASLSPASTPSCLSAASSIDNLG 997
QY 781 SFSELSVVSSSGTSGASLEKK--EVPQVDFPSITQFVRNLGLEHMDIFEREQITLDVL 838
Db 998 PLAELAVGASNAGDAAGTERKEGEVAGLDNISOFLKSLGLEHLRIDIFETEQITLDVL 1057
QY 839 VEMGHKELKEIGINAYGHRHKLIGVERLISQOGLNPYLTNTSGSGTILIDLSPODKE 898
Db 1058 ADMGHEELKEIGINAYGHRHKLIGVERLISQOGLNPYLTNPHFCVNOGTILIDLAPEDKE 1117
QY 899 FOSVEEMOSTVREHRDGGHAGGINFRYNNILKIQVCKNKLWERYTHRRKEVSEENHHA 958
Db 1118 YOSVEEMOSTVREHRDGGHAGGINFRYNNIRIQVCKNKLWERYTHRRKEVSEENHHA 1177
QY 959 NERMPLFGSPFNATITHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYGGGTCPCPVH 1018
Db 1178 NERMPLFGSPFNATITHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYGGGTCPCPTH 1237
QY 1019 KDRSCYICHRQLLCFRTVLGKSLFLOFSAMKMAHSPGHHSTVGRPSVNGALAEVYVIRG 1078
Db 1238 KDRSCYICHRQLLCFRTVLGKSLFLOFSAMKMAHSPGHHSTVGRPSVNGALAEVYVIRG 1297
QY 1079 EQAYPEYLITYQIMRPE 1095
Db 1298 EQAYPEYLITYQIMRPE 1314

RESULT 3

US-09-196-387-10

; Sequence 10, Application US/09196387

; Patent No. 6277613

; GENERAL INFORMATION:

; APPLICANT: de Lange, Titia

; APPLICANT: Smith, Susan

; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS

; TITLE OF INVENTION: OF USE THEREOF

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber & Jackson

; STREET: 411 Hackensack Avenue, 4th Floor

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/196,387

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/095,225

; FILING DATE: June 10, 1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 949 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-196-387-10

Query Match 53.6%; Score 3093; DB 4; Length 949;

Best Local Similarity 81.5%; Pred. No. 2e-272;

Matches 585; Conservative 60; Mismatches 71; Indels 2; Gaps 1;

QY 1 GFGKDVVEYLLQNGASVQARDGGGLPLHNACSFHGAHVNVNLLRHGADPNARDNNYT 60
Db 225 GFGKDVVEYLLQNGASVQARDGGGLPLHNACSFHGAHVNVNLLRHGADPNARDNNYT 284
QY 61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKKDELLESAR 120
Db 285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGSALDADPSAKAVLTGEYKKDELLEAAR 344
QY 121 SGNEEKWALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVNADKGGDLV 180
Db 345 SGNEEKWALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVNADKGGDLV 404
QY 181 PLHNACSYGHVEYTELLVKGACVNMADLMQFTPLHEAAAKNRVEVCSLLLSYGADPTLL 240
Db 405 PLHNACSYGHVEYTELLVKGACVNMADLMQFTPLHEAAAKNRVEVCSLLLSYGADPTLV 464
QY 241 NCHNKAIDLAPTOLKELAYEFKSHLSLOAREADVTRIKKLSLEMVNFKHPQTHET 300
Db 465 NCHGKSAVDMAPTPELRERTYEFKSHLSLOAREADLAKYKKTALAEIINFKQPOSHET 524
QY 301 ALHCAASPYKPKOICELLKRGANINEKTEFTPLHVASEKAHNDVVEVVKHEAKV 360
Db 525 ALHCAVASLHPKRVKOVTELLLRKANVNEKNKDEMTPLHVAARAHNDVMEVLKHKHAKM 584
QY 361 NALDNLGQTSLHRAAYCGHLOTCLRLILSYGCDPNIIISLQGTALQMGNNVQQLQESIS 420
Db 585 NALDTLGTALHRAALAGHLOTCLRLILSYGSDPSIISLQGTAAQMGNEAVQQLSESTP 644
QY 421 LGNSEADQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVWEYLL 480
Db 645 IRTSDVYRLLLEASKAGDLETVKOLCSSQNVNCRDIEGRHSTPLHFAAGYNRVSVWEYLL 704
QY 481 OHGADVHAKDGGGLVPLHNACSYGHEVAELLVKGAVNVNADLWKFTPLHEAAAKGKYE 540
Db 705 HHGADVHAKDGGGLVPLHNACSYGHEVAELLVKGAVNVNADLWKFTPLHEAAAKGKYE 764
QY 541 ICKLLQHGADPTKKNRGNTPLDLVKDGDTDIDQLLRGDAALLDAKKGCLARVKKLSS 600
Db 765 ICKLLKHGADPTKKNRGNTPLDLVKDGDTDIDQLLRGDAALLDAKKGCLARVKKLCT 824
QY 601 PDVNCRTDQGRHSTPLHLAAGYNLEVAEYLLQHGADVNADKGGGLPLHNAASYGH-- 658
Db 825 PENINCRTDQGRNSTPLHLAAGYNLEVAEYLLHGHGADVNADKGGGLPLHNAASYGCL 884
QY 659 VDAVALLIKYNACVNTADKWAFTPLHEAAQKGRTOCALLLAHGADPTLKNQEGQTP 716
Db 885 ARVQKLCPTENINCRTDQGRNSTPLHLAAGYNLEVAEYLLHGHGADVNADKGGGLPL 942

RESULT 4

US-09-196-387-8

; Sequence 8, Application US/09196387

; Patent No. 6277613

; GENERAL INFORMATION:

APPLICANT: de Lange, Titia
APPLICANT: Smith, Susan
TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
TITLE OF INVENTION: OF USE THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/196,387
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/095,225
FILING DATE: June 10, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 673 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-196-387-8

Query Match 33.3%; Score 1918.5; DB 4; Length 673;
Best Local Similarity 80.9%; Pred. No. 1.1e-165;
Matches 364; Conservative 42; Mismatches 43; Indels 1; Gaps 1;
QY 1 GFRKDVVEYLLQNGASVQARDGGGLPLHNACSFSGHAEVYVNLRLRHGADPNARDNNY 60
DB 225 GFRKDVVEYLLQNGASVQARDGGGLPLHNACSFSGHAEVYVNLRLRHGADPNARDNNY 284
QY 61 PLHEAAIKGKIDYCVILLOHGAEPITRNTDGRALDLPDSAKAVLTGEYKKDELLESA 120
DB 285 PLHEAAIKGKIDYCVILLOHGAEPITRNTDGRALDLPDSAKAVLTGEYKKDELLESA 344
QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLOHGAADVHAKDKGLV 180
DB 345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLOHGAADVHAKDKGLV 404
QY 181 PLHNACSYGHEVTELLVKGACVNDMLQFPTPLHEAASKNVEVCSLLLSYGADPTLL 240
DB 405 PLHNACSYGHEVTELLVKGACVNDMLQFPTPLHEAASKNVEVCSLLLSYGADPTLL 464
QY 241 NCHNSAIDLAPTOLKERLAYEFKGSLLQAAAREADVTRIKHLSLEWYVNFKHQTHET 300
DB 465 NCHGSAVDMAPTPELRUTYEFKGSLLQAAAREADVTRIKHLSLEWYVNFKHQTHET 524
QY 301 ALHCAASPYPKRKOICELLRRKGANINEKTEFLTPHVAEKAHNDVVEVVKHEAKV 360
DB 525 ALHCAVASLHPKQVTELLRRKGANINEKTEFLTPHVAEKAHNDVVEVVKHEAKV 584
QY 361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDDNIIISLQGTALQNGENYVQQLLEGIS 420
DB 585 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDDNIIISLQGTALQNGENYVQQLLEGIS 644
QY 421 LGNSEADRLLEAAKAGDVETVKKLCTVQS 450

DB 645 SDPSIISIQGTAAQMGNEAVQOILSGHS 673

RESULT 5

US-09-172-977-4
Sequence 4, Application US/09172977
Patent No. 5989863
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Yue, Henry
TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
FILE REFERENCE: PF-0615 US
CURRENT APPLICATION NUMBER: US/09/172,977
CURRENT FILING DATE: 1998-10-14
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PERL Program
SEQ ID NO 4
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: g29491
US-09-172-977-4

Query Match 14.9%; Score 860.5; DB 2; Length 1839;
Best Local Similarity 29.1%; Pred. No. 1.5e-68;
Matches 297; Conservative 147; Mismatches 404; Indels 174; Gaps 29;

QY 3 GRKDVVEYLLQNGASVQARDGGGLPLHNACSFSGHAEVYVNLRLRHGADPNARDNNY 62
DB 75 GHVGLVQELLRGSSVDSATKKNTALHSLAGQAEVYVNLREGANINQAQSQNGETPL 134
QY 63 HEAAIKGKIDYCVILLOHGAEPITRNTDGRALDLPDSAKAVLTGEYKKDELLESA 115
DB 135 YMAAQENHIDVVKYLLLENGANQSTATEDGGTPLAVALOOGHNOAVILLENDTKGKRLP 194
QY 116 LESARSGNEEKKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQL 163
DB 195 ALHIAARKDDTKSAALLQ---NDHNADVQSKMMVNRRTTESGFTPLHIAAHYGVNVA 251
QY 164 LLOHGAADVHAKDKGLVPLHNACSYGHEVTELLVKGACVNDMLQFPTPLHEAASKN 223
DB 252 LLNREGAADVDTARNGITPLHVASKRGNTVMKLLDRGGQIDAKTRDGLTPLHCAARSG 311
QY 224 VECSLLLSYGADPTLLNCHNSAIDLAPTOLKERLAYEFKGSLLQAAAREADVTRIK 283
DB 312 DVVVELLLERGA-PLIARTKN-----GLSPLHMAAQGDHVECVK 349
QY 284 HLSLEWYVNFKH-----QTHETALHCAASPYPKRKOICELLRRKGANINEKTEFL 338
DB 350 HL-----LQHKAPVDVDTLDYLTALHVA---HCGHYEVTKLLDKRANPARALNGET 402
QY 339 HVASEKAHNDVVEVVKHEAKVNDMLGQTSLHRAAYCGHLQTCRLLLSYGCDDNIIIS 398
DB 403 HIACKNRIKVMELLVKGASIQAITESGLTPIHVAAFMGHLNIVLLLLONGASPDVTNI 462
QY 399 QGTALQNGENYVQQLLEGISLGNSEADRLLEAAKAGDVETVKKLCTVQSVNCRDIEG 458
DB 463 RGETALHM-----AARAGQVEVVR--CLLRNGALVDARA 494
QY 459 R-OSTPLHFAAGYNRVSVVEYLLLOHGAADVHAKDKGLVPLHNACSYGHEVVAELLVKH 517
DB 495 REQOTPLHITASRLGKTEIVOLLLOHMAHPDAATTNGYTPHLHISAREQVDVASVLF 554
QY 518 VNVADLWFTPLHEAAKAGYETCKLLLOHGAADVTKKNDGNTPLDVLVKGDTDTODLL 577
DB 555 AHSLATKKGFTPLHVAAYGSLDVAKLLLORRAAADSAGKNGLTPLHVAHYDN----- 608
QY 578 RGDAAILLDAKKGCLARVKLSSPDVNVNCRDGTGRHSTPLHLAAGYNNLEVAEYLLQ 637

Db 609 -----OKVALLLEKASP-----HATAKNGYTPPLHIAAKKNQMOIASTLLNYGA 653
QY 638 DYNADQKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTOALCAL 697
Db 654 ETNIVTKOGVTPPLHLSAQEGHTDMVTLLLDKGANIHMTSGSLTSLHAAQEDKVNVDI 713
QY 698 LLAHAGDPTLKNQOGTDLIV-----SADDVSALLTAAMPSPA-LPSCYKP----- 743
Db 714 LTKHAGADAHKGLYTPPLIVACHYGNVKNVFLKQGANVNAKTKNGYTPPLHQAQOQH 773
QY 744 -OVLNGVRSPGATADALSSGSPSSLSAASSLNLGSGFSELSSVSSSGTEGASLEK 802
Db 774 THIIINVLLQHGAKPNATTA--NGNTALAIARLGYI--SVVDTLKVVVTEETTTTTITE 829
QY 803 KEVPGVDFSTOFVNLGLEHLMDFERE---QITLDVLVEMGHKELKEIGINAYGHRHK 859
Db 830 KHLNVPETMTE-----VLVDSEEGDDTMTGOGGEYLRPEDLKELGDDS----- 874
QY 860 LITKVERLISGO--OGLNPYLTLNTSGSTILIDLSPDKKEFQSVSEEMQSTVREHRDGG 917
Db 875 -----LPSSQFLDGMN-YLRYSLGEG-----RSDSLRSFSDRSHTLSHASYLRD-- 918
QY 918 HAGGFNRYNLIKQKVCN-KKLMBRYTHRRKEVSEENHNHANERMLPHGSPFVNAIHK 976
Db 919 --SAYMDDSVIPSHQVSTLAKAERNRYLSWGTENLDNVA-----LSSSP-----IHS 966
QY 977 GF 978
Db 967 GF 968

RESULT 6

US-09-172-977-3

; Sequence 3, Application US/09172977

; Patent No. 5989863

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Yue, Henry

; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN

; FILE REFERENCE: PF-0615 US

; CURRENT APPLICATION NUMBER: US/09/172.977

; CURRENT FILING DATE: 1998-10-14

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PERL Program

; SEQ ID NO 3

; LENGTH: 843

; TYPE: PRT

; ORGANISM: Rattus norvegicus

; FEATURE: -

; OTHER INFORMATION: g1841966

US-09-172-977-3

Query Match 14.6%; Score 841.5; DB 2; Length 843;
Best Local Similarity 29.7%; Pred. No. 2.1e-67;
Matches 265; Conservative 138; Mismatches 351; Indels 137; Gaps 21;

QY 3 GRKDVVEYLLONGASVQARDGGLIPLHNAASFGHAEVYNLLLRHGADPNARDNNYTP 62
Db 40 GHVGLVQELLGRGSSVDSATKGNLTALHIASLAGQAEVVKVILVKEGANINAOQSGFTPL 99
QY 63 HEAAIKGKIDVCIVLLQHGAEPTIRMTDGR-----ALDLADPSAKAVLTGEYKKDEL--- 115
Db 100 YMAAQENHIDVVKYLLLENGANQSTATEDGFTPLAVALQOQHQAVALLENDTKGVRLP 159
QY 116 -LESARGNEEKMWALLTPLNVNCHASDGRKS-----TPHLAAGYNRVKIVOL 163
Db 160 ALHIAARKDDTKSAALLQ---NDHNADVQSKMWNRTTESGFTPLHIAHYGNVNVATL 216
QY 164 LLOHGADVHAKDGLDVLPLHNAASYGHEVTELLVKGACVNAMDLWQFTPLHEAASKNR 223
Db 217 LLNRGAADVFTARGITPLHVASKRGNTNMVKLLLDGRGOIDAKTRDGLTPLHCAARSGH 276

QY 224 VEYCSLLSYGADPTLLNCHNKSALDLAPTQPKERLAYEFKGHSLLOAARADVYTRIKK 283
Db 277 DQVVELLLERGA--PLLARTKN-----GLSPHMAAQGDHVECVK 314
QY 284 HLSLEMVNFKHP-----QTHETALHCAASPYPKRQICICELLKRCANINEKTEPLTPL 338
Db 315 HL-----LOHKAPVDDVTLDYLTALHVA--HCGHYRVTKLLLDKRANPNARALNGFTPL 367
QY 339 HVASEKAHNDVVEVYVYKHEAKVNALDNLGQTSLHRAAYGCHLQTCRLLSYSGDPNIISL 398
Db 368 HIACKNRKRYKMBELLYKAYIQAITESGLTPIPVAAPFMGHLNIVLLLLQNGASPDVTNI 427
QY 399 QGFTALOMGNENVOQLLOEGISLGNSEADQRLLEAAKAGDVETVKKLTQVSVNCRDIEG 458
Db 428 RGETALHM-----AARAGEVEVVR--CLLRNGALVDARA 459
QY 459 R-OSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGGLVPLHNAASYGHEVAELLVKHGA 517
Db 460 REQOTPLHIASRLGKTEIVOLLQHMAHPDAATNGTTPPLHISAREGQVDVASVLEAGA 519
QY 518 VNVADLWKFTPLHEAAAKGYEICKILLQHGADPTKKNRDGNTPLDLVKDGDIDQLL 577
Db 520 AHSLATKKGFTPLHVAKYGSLDVAKLLLORRAADSAGKNGLTPLHVAHYDN----- 573
QY 578 RGDAAALDAAKGCLARVKKLSSPDNVNCRDGTGRHSTPLHLAGYNNLEVAEYLLQHGA 637
Db 574 -----QKVALLLLEKASP-----HATAKNGYTPPLHIAAKKNQMOIASTLLNYGA 618
QY 638 DVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTOALCAL 697
Db 619 ETNTVTKOGVTPPLHLSAQEGHTDMVTLLVLEKGANIHMTSGSLTSLHAAEDKVNVDI 678
QY 698 LLAHAGDPTLKNQOGTDLIV-----SADDVSALLTAAMPSPA-LPSCYKP----- 743
Db 679 LTKHAGADAHKGLYTPPLIVACHYGNVKNVFLKQGANVNAKTKNGYTPPLHQAQOQH 738
QY 744 -OVLNGVRSPGATADALSSGSPSSLSAASSLNLGSGFSELSSVSSSGTEGASLEK 802
Db 739 THIIINVLLQHGAKPNATTA--NGNTALAIARLGYI--SVVDTLKVVVTEETTTTTITE 794
QY 803 KEVPGVDFSTOFVNLGLEHLMDFERE---QITLDVLVEMGHKELKEIG 850
Db 795 KHLNAPETMTE-----VLVDSEEGDDTMTGOGGEYLRPEDLKELG 836

RESULT 7

US-09-031-485-33

; Sequence 33, Application US/09031485

; Patent No. 5824306

; GENERAL INFORMATION:

; APPLICANT: Tang, Liang

; APPLICANT: Blehm, E. Scot

; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND

; TITLE OF INVENTION: USES THEREOF

; NUMBER OF SEQUENCES: 85

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carol Talkington Verser, Ph.D.

; ADDRESS: Heskia Corporation

; STREET: 1825 Sharp Point Drive

; CITY: Fort Collins

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80525

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Windows 95

; SOFTWARE: WordPerfect for Windows, Version 7.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/031.485

; FILING DATE:

```

; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-485-33

Query Match      14.2%; Score 821; DB 2; Length 1745;
Best Local Similarity 29.1%; Pred. No. 5.6e-65;
Matches 276; Conservative 125; Mismatches 321; Indels 228; Gaps 28:

QY   3 GRKDVEVLLONGASVQARDGGGLIPLHNACSFGEAEVVNLLLRHGADPNARNWNYTPL 62
    | | | : | : | : | : | : | | | | | | | | | | | | | | | | | | |
Db   48 GNLDVLELLRSRGTDINTCNANGLNALHLSKEGHEHVRELLKRKADVDAATRKGNAL 107

QY   63 HEAAIKRIDVICVILLOHGAETPIRTDGTALDLADP-----SAKAVLTGE 109
    | | | : | : | : | : | : | | | | | | | | | | | | | | | | | | |
Db   108 HIASLAGQLIVTLVENGANVNQSNGFTPLYMAAENHESVRYLLAHNAQAALSTE 167

QY   110 YKDELLESARSGNEEKMMALLTPLNVNCHASDGKRSTP-LHLAAGYNRVKIVQLLQ-- 166
    | | | : | : | : | : | : | | | | | | | | | | | | | | | | | | |
Db   168 DGFPTPLAVALQQOHRVAVVLE-----NDTRGVRLPALHIAAKKDDTKRAATILLQNE 221

QY   167 HGADVHAOKDGLVPLHNACSXYHYETELLVKHGACVNAMDLNQFTPLPHFAASKNRVEV 226
    | | | : | : | : | : | : | | | | | | | | | | | | | | | | | | |
Db   222 HNSDV--TSKSGTTPHIAAHGNNVAQLLEKGANVYQARHNI SPLHVATKWGRTNM 279

QY   227 CSLLSYGDPTLLCNCHKNSAIDLAPTQLKERLAYEFKGHSLLQAAREADVTRIKKHL 286
    ||||| : | : | : | : | : | | | | | | | | | | | | | | | | | |
Db   280 VSULLAHGA---VIDCKTRDLL-----298

QY   287 LEMVNFKHQTHETALHCAASPYPKRKQICEILLRGANINEKTKEFLTPLHVASEKAH 346
    | | | : | : | : | : | : | | | | | | | | | | | | | | | | | | |
Db   299 -----TPLHCSRSRH---DQVDVLELLEGAPISAKTKNGLAPLHMAAQ--- 339

QY   347 NDVEVYVHKAEKVNALDNLGOSTLHPAAAYCGHLOTCLRLLSYCCDPNIISLOGFTALQM 406
    | | | : | : | : | : | : | | | | | | | | | | | | | | | | | | |
Db   340 --VDDTVDY-----LTPLVAAHCGHVRAKLLEDNRADPNARALNGFTPLHI 386

QY   407 G-----NENVQOLLQEGIS-----LGNSEADROLLEAAKAGDVEIVKK 444
    : : : : | : | : | : | : | | | | | | | | | | | | | | | | | | |
Db   387 ACKKNRIKIVELLLKVHAAIEATESGLSPLHVAFPGANIVYLLQQCANADVAIV-- 444

QY   445 LCTVQSVNCRDEGROSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGGLVPLHNACSYG 504
    | | | : | : | : | : | : | | | | | | | | | | | | | | | | | | |
Db   445 -----RGTEPLHAAARANQTDIRVLVRNGAQVDAAAARELQTPHLHIA SRLG 490

QY   505 HYEVAELLVKHAVNVNADLWKFTPLHEAAAKGYEICKLLLOHGADPTKKNRDGNTPLD 564
    | | | : | : | : | : | : | | | | | | | | | | | | | | | | | | |
Db   491 NTDIVILLQANASPNAATRDLYPLHIAAKEGQEEVAAILMDHGTDKTLTKKGETPLH 550

QY   565 LVKDGDTDIQDLRLRGDAALLDAAKGCLARVKKL-----SSPDNVNCRDTCGRHS-TPLHLA 620
    | | | : | : | : | : | : | | | | | | | | | | | | | | | | | | |
Db   551 L-----AAKYGNLPVAKSLERGP-----VDIEGKNQVTPLHVA 585

QY   621 AGYNNLEVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAF 680
    | | | : | : | : | : | : | | | | | | | | | | | | | | | | | | |
Db   586 AHYNDKVALLLLENGASAHAANKGYTPLHIAAKKNQMDIASTLLHYKANANAESKAGF 645

QY   681 TPLHEAAOKGRTOLCALLLAHAGADPTLKNOEGOTPLDLYSADDVSALL-----TAMP 734
    ||||| ||| : | : | : | : | : | | | | | | | | | | | | | | | | | |

```

[illegible]

Db 586 AHYNNDKVALLENGASAHAAKNGYTPHLIAAKNQMDIASTLLHYKANANAESKAGF 645
 QY 681 TPLHEAAQKGRTOCALLLAHGADPTLKNQEGOTPLDLVSADDSALL-----TAAMP 734
 Db 646 TPLHAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCQAQEDRVSAEELVKENAAIDP 705
 QY 735 SALP-----SCYKPOVLNGVR---SPGATADALSSGSPSSLSAASSLDNLGSGFSE 784
 Db 706 KTKAGYTPLVACHFGQI-NMVRFLIEHGARGVSVITRASYTPHLHQAQOG----- 754
 QY 785 LSSVSSSGTEGASLEKKEVPGVDFSIQ-----FVRNLG-----LEHMDIFEREQIT-L 835
 Db 755 HNSVRYLLEHGAS-----PNVHTSGOTPLSLAEELGVSVVEALKITITETVTITET 807
 QY 836 DVLVEMGHK-----ELKEIGINAYGH--RHKLIKGV 864
 Db 808 TTVTEERYKQPONPEAMNETMFSDEDEGDNQITANAHADFSESLTKGL 857

RESULT 10

US-09-557-034-33

; Sequence 33, Application US/09557034

; Patent No. 6365569

; GENERAL INFORMATION:

; APPLICANT: Tang, Liang

; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN

; PROTEINS, NUCLEIC ACID MOLECULES, AND

; USES THEREOF

; NUMBER OF SEQUENCES: 171

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carol Talkington Verser, Ph.D.

; STREET: 1825 Sharp Point Drive

; CITY: Fort Collins

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80525

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Windows 95

; SOFTWARE: WordPerfect for Windows, Version 7.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/557,034

; FILING DATE: 21-Apr-2000

; CLASSIFICATION: <Unknown>

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 09/065,474

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Verser, Carol Talkington

; REGISTRATION NUMBER: 37,459

; REFERENCE/DOCKET NUMBER: HW-5-C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 970/493-7272

; TELEFAX: 970/484-9505

; INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1745 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 33:

US-09-557-034-33

Query Match

Best Local Similarity 14.2%; Score 821; DB 4; Length 1745;

Matches 276; Conservative 125; Mismatches 321; Indels 228; Gaps 28;

QY 3 GRKDVVEYLQNGASVQARDGGLPLHNAACSFGRHAEVVNLLLRGADPNADNNMYTPL 62

Db 48 GNLDVRVLELLRSRGTINTCNANGLNALHLASKEGHHVVRELLKRKADYDAATRKGNTAL 107
 QY 63 HEAAIKGKIDVICVILQHGAEPTIRNTDGTALDLADP-----SAKAVLTGE 109
 Db 108 HIASLAGQELIVTLVENGANVNVQSLNGFTPLYMAAQENHESVVRVYLLAHNANOALSTE 167
 QY 110 YKKDELLSARSNGEEMKWMALLTPLNVNCHASDGKRSKP-LHLAAGYNRVKIVQLLLQ-- 166
 Db 168 DQFTPLVALQOQHDRVAVVLE-----NDTRGVRLPALHIAAKKDDTKAATLLQNE 221
 QY 167 HGADVHAKDKGDLVPLHNACSYGHVTELLVKHGACVNMADLQWOTPLHEAASKNRREV 226
 Db 222 HNSDV--TSKSGFTPLHAAHGNVENVAQLLEKGANVYQARHNISPLHVATKWRGNTM 279
 QY 227 CSLLSYGADPTLLNCHNKSAIDLAPTLPOLKERLAYEFKGHSLLOAAREADVTRIKKHL 286
 Db 280 VSLLLAHGA--VIDCTRDL-- 298
 QY 287 LEMVFNKHPQTHETALHCAASPYPKKQICELLRLKGANINEKTEFELPLHVASEKAH 346
 Db 299 -----TPLHCASRSGH---DQVVDLLEKGAIPISAKTKGLAPLHMAAQ--- 339
 QY 347 NDVVEVVVHKAQVNAIDNLGQTSLHRAAYCGHLOTCLLLSYGCDPNIISLOGFTALQM 406
 Db 340 --VDDVTVDY-----LTPLHVAACHGVRAKLLLDNRADPNARALNGFTPLHI 386
 QY 407 G-----NENVOOLLQEGIS-----LONSEADROLLPEAAKAGDVETVKK 444
 Db 387 ACKNRIKIVELLKYLHAAIEATTESGLSPHVAAFMGAINIVLYLLOOGANADVTV-- 444
 QY 445 LCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVYLLQHGADYHAKDKGGLVPLHNACSYG 504
 Db 445 -----RGETPLHAAARANTDILRVLRNQAQVDAARELOTPPLHIASRLG 490
 QY 505 HYEVAELLVKHGVNVADLWKFTPLHEAAKGYEICKLLLOHGADPTKKNRDGNTPLD 564
 Db 491 NTDIVILLQANASPNAAATRDLYTPHIAAKEGQEEVAAIILMDHGTDKTLTKKGFTPLH 550
 QY 565 LVKDGDTDIODLLRGDAALLDAKKGCLARVKL---SSPDNVNCRDTQGRHS--TPLHLA 620
 Db 551 L-----AAKYNLPVAKSLERGT-----VDIEGKNQVTPHLVA 585
 QY 621 AGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVVAALLIKYANCVNATDKWAF 680
 Db 586 AHYNNDKVALLENGASAHAAKNGYTPHLIAAKNQMDIASTLLHYKANANAESKAGF 645
 QY 681 TPLHEAAQKGRTOCALLLAHGADPTLKNQEGOTPLDLVSADDSALL-----TAAMP 734
 Db 646 TPLHAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCQAQEDRVSAEELVKENAAIDP 705
 QY 735 SALP-----SCYKPOVLNGVR---SPGATADALSSGSPSSLSAASSLDNLGSGFSE 784
 Db 706 KTKAGYTPLVACHFGQI-NMVRFLIEHGARGVSVITRASYTPHLHQAQOG----- 754
 QY 785 LSSVSSSGTEGASLEKKEVPGVDFSIQ-----FVRNLG-----LEHMDIFEREQIT-L 835
 Db 755 HNSVRYLLEHGAS-----PNVHTSGOTPLSLAEELGVSVVEALKITITETVTITET 807
 QY 836 DVLVEMGHK-----ELKEIGINAYGH--RHKLIKGV 864
 Db 808 TTVTEERYKQPONPEAMNETMFSDEDEGDNQITANAHADFSESLTKGL 857

RESULT 11

US-09-082-059-2

; Sequence 2, Application US/09082059A

; Patent No. 6225086

; GENERAL INFORMATION:

; APPLICANT: Morrow, Jon S.

; APPLICANT: Devorajan, Prasad

; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identifi

; FILE REFERENCE: 44574-5002-US

; CURRENT APPLICATION NUMBER: US/09/082,059A

wed Feb 12 16:08:51 2003

```

; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: 60/047356
; EARLIER FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-082-059-2

```

```

Query Match      10.9%   Score 629.5;   DB 4;   Length 1088;
Best Local Similarity 28.1%;   Pred. No. 6.7e-48;
Matches 195;   Conservative 85;   Mismatches 234;   Indels 181;   Gaps 16;

Qy 164 LQHGADVHARDKGDVPLPHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNR 223
Db 7 LLDKKNPNKALNGFTPLHACKNRIKVMELLKHGASIQAVTESGLTPIHVAAPMGH 66

Qy 224 VEVCSLLSYGADPTLLNCHNKSALIDAPTPQLKRLAYEFKHSILQAAREADVTRIKK 283
Db 67 VNVISQLMHGASPTTNRVG----- 87

Qy 284 HLSLEWVFKHPQHTETALHCAASAPYKPKKOICELLRKGANINEKTEFLTPHVAASE 343
Db 88 -----ETALHMAARS---GQAEVYRLVDOGAQVEAKAKODQTPHISAR 129

Qy 344 KAHNDVVEVVKHEAKVNALDNGQTSUHRAYCGHLOTCRLLSYGCDPNIIISLOQFETA 403
Db 130 LQKADIVQOLLQOGASPNAAATTSYTPHLHLSAREGHEDVAAFLLDHGASLSITTKKGF-- 187

Qy 404 LQMGENVQOLLQEGISLGNSEADRLLEAAKAGDVETVKLCITVQSVNCRDIEGRQSTP 463
Db 188 -----TP 189

Qy 464 LHFAAGYNRVSVVEYLQHGADVHAKDGLVPLHNACSYGHEVYAEILLVKGAVNVVAD 523
Db 190 LHVAAYKYLEVANLLLOKSASPDRAKSGSLTPLHVAHYDNQKVALLLDQOGASPHAAA 249

Qy 524 LMKFTPLHEAAKGYEICKLLQHGADPTKKNRDNPTDL-VKDGDTDTQDILLRGDAA 582
Db 250 KNGYTPHIAAKKNQMDIATTLLEYGADANAVTROGIAVSHLAAQEGHVMVSVLLGRNA 309

Qy 583 LLDAAKKGCLARVKLLSPDNVNCRTQCRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQ 642
Db 310 -----NVNLSNKG--LTPHLAAQEDRVNVAEVLNQGAVHDAQ 347

Qy 643 DKGGGLPLHNAASYGHDVVAALLIKYNACVNATDKWAFPLHEAAKQRTQCALLLAHG 702
Db 348 TKMGYTPPLHVGHYGNIKIVNELLQHSKAVNAKNGYTPHQAQQGQHTHIIYNLLQNN 407

Qy 703 ADPTLNQEGQTPDL-----VSADDYSALLT-AAMPSPALPSYCK---POVLNGV--R 750
Db 408 ASPNELTVNGNTALGIARBLGYISVVDTLKIIVTEMTTTTTTEKHKNVPEETMEVLDM 467

Qy 751 SPQATADALSS-----GP-----SSPS-----SLSAASLDNLGSGFSLSVV 789
Db 468 SDEGEDAMTQDVKYLGQDLKELGDDSLPAFGYMGFSLGARSASDR---SY-----TLN 520

Qy 790 SSSGTEGASLEKVEYGVDFSTQTVRNIGLEHL 824
Db 521 RSSYARDSMMIEELLVPSKEQHLT-FTREFDSDSL 554

```

```

RESULT 12
US-09-065-474-139
; Sequence 139, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND

```

```

; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Hesk Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-09-065-474-139

```

Query Match 8.9%; Score 513; DB 3; Length 352;

Best Local Similarity 30.2%; Pred. No. 3.7e-38; Indels 86; Gaps 5;

Matches 127; Conservative 64; Mismatches 144; Indels 86; Gaps 5;

```

Qy 299 ETALHCAASAPYKPKKOICELLRKGANINEKTEFLTPHVAASEKAHNDVVEVVKHEA 358
Db 12 ETPLHLAARA---NQTDIVRVLVNAGVDAARELOTPHLIASRLGNTDIVILLQANA 68

Qy 359 KVNALDNLGQTSUHRAYCGHLOTCRLLSYGCDPNIIISLOQFTALQMGENVQOLLQEG 418
Db 69 SPNAATRDLYTPHLIAAKEGQEEVAAILMDHGTDKTLTKKGFPLHL----- 116

Qy 419 ISLGNSEADRLLEAAKAGDVETVKLCITVQSVNCRDIEGR-QSTPLHFAAGYNRVSVVE 477
Db 117 -----AAKYNLPVAKSL--LERGTPVDIEGKNQVTPHVAHYNNDKVAL 160

Qy 478 YLQHGADVHAKDGLVPLHNACSYGHEVYAEILLVKGAVNVVADLWKFTPLHEAAK 537
Db 161 LLENGASAHAAKNGYTPHLIAAKKNQMDIATLLHYKANANAEKAGETPLHLAAOEG 220

Qy 538 KYEICKLLQHGADPTKKNRDNPTDLVLVDGDDTDIODLLRGDAALLDAKKGCLARVK 597
Db 221 HREMAALLIENGAKVGAQARNG----- 242

Qy 598 LSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYG 657
Db 243 -----LTPMHLCAQEDRVSAEELVKENAAIDPKTKAGYTPHVAACHFG 286

Qy 658 HVDVAALLIKYNACVNATDKWAFPLHEAAKQRTQCALLLAHGADPTLNQEGQTPDL 717
Db 287 QINMVRFLIEHGAARVSVITRASYPHQAQQGQHSVVYRLLLEHGAASPNVHTSTGTPLS 346

Qy 718 L 718
Db 347 I 347

```

```

RESULT 13
US-09-557-034-139

```


Db 118 NOTDIVRVLRNGAQVDAARELOTPHLIASRLGNTDIVILLQANASPNAAATRDLYTPL 177
Qy 564 DL-VKDGTDIDODLLRG---DAALLDAAKKGLARVKLLSSPDNVNCRDQTGRHSTPLHL 619
Db 178 HIAAKEGQEEVAAILMDHGTDKTL--TKKG-----FTPLHL 212
Qy 620 AAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACYNATDKWA 679
Db 213 AAKYGNLPVAKSLLEGTPTVDIEGKNQVTPLVAAHYNNNDKVALLENGASAHAAKNG 272
Qy 680 FTPLHEAAOKGRTOLCALLLAHAGDPTLKNQ 710
Db 273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303

RESULT 15

US-08-847-429A-23
; Sequence 23, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scott
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-847-429A-23

Query Match 7.3%; Score 421.5; DB 2; Length 303;
Best Local Similarity 30.4%; Pred. No. 6.2e-30;
Matches 119; Conservative 52; Mismatches 125; Indels 95; Gaps 8;
Qy 327 INEKTKEFLPLHVAASEKAHNDVVEVVVVKHEAKVNALDNLGQTSLHRAAYCGHLOTCRL 386
Db 1 VDDVTVDYLTPLHV-----AAHCGHVRVAKLL 27
Qy 387 LSYGCDPNIISLOGFTALQMNENVOQLQEGISLGNSEADROLLEAAKAGDVETVKKLC 446
Db 28 LDRNADPNARALNGFTPLHI-----ACKKNRIKIVLELL 61
Qy 447 TVQSVNCRDIEGRQS---TTLPHFAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNACSY 503
Db 62 KYHAA-----IETTTESGLSLPHVAAMGAINIVYLLQOGANADVATVRGETPLHLAARA 117

Qy 504 GHYEVAELLVKHGA VNVADLWKFTPLHEAAAKGYEICKLLLOHGADPTKKNRDNTPL 563
Db 118 NOTDIVRVLRNGAQVDAARELOTPHLIASRLGNTDIVILLQANASPNAAATRDLYTPL 177
Qy 564 DL-VKDGTDIDODLLRG---DAALLDAAKKGLARVKLLSSPDNVNCRDQTGRHSTPLHL 619
Db 178 HIAAKEGQEEVAAILMDHGTDKTL--TKKG-----FTPLHL 212
Qy 620 AAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACYNATDKWA 679
Db 213 AAKYGNLPVAKSLLEGTPTVDIEGKNQVTPLVAAHYNNNDKVALLENGASAHAAKNG 272
Qy 680 FTPLHEAAOKGRTOLCALLLAHAGDPTLKNQ 710
Db 273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303
Search completed: February 12, 2003, 12:27:18
Job time : 21.0427 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 03:33:09 ; Search time 22.5641 Seconds
(without alignments)
1245.508 Million cell updates/sec

Title: US-09-843-159b-3

Perfect score: 5769

Sequence: 1 GFRKDDVVEYLLQNGASVQA.....AYPEYLITYQIMRPGMWDG 1100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5602	97.1	1074	10	US-09-509-196A-2
2	4889.5	84.8	1327	10	US-09-841-835-2
3	3596	62.3	802	9	US-09-964-899-41
4	3093	53.6	949	10	US-09-841-835-10
5	1918.5	33.3	673	10	US-09-841-835-8
6	798.5	13.8	1724	9	US-09-964-899-43
7	395	6.8	426	10	US-09-908-711-70
8	379.5	6.6	1719	9	US-09-012-896-378
9	379.5	6.6	1719	9	US-09-895-793-378
10	379.5	6.6	1719	9	US-09-895-814-378
11	379.5	6.6	1719	10	US-09-759-143-378
12	379.5	6.6	1719	10	US-09-780-669-378
13	379.5	6.6	1719	10	US-09-822-827-378
14	368.5	6.4	740	10	US-09-835-788A-12
15	332.5	5.8	835	10	US-09-947-199-2
16	324.5	5.6	835	10	US-09-947-199-8
17	312.5	5.4	599	10	US-09-735-388-2
18	297	5.1	551	10	US-09-835-788A-17
19	281.5	4.9	285	10	US-09-835-788A-18

20	268	4.6	306	10	US-09-835-788A-19	Sequence 19, Appl
21	267	4.6	1054	10	US-09-798-042-87	Sequence 87, Appl
22	258	4.5	251	10	US-09-835-788A-13	Sequence 13, Appl
23	247.5	4.3	328	10	US-09-758-593A-11	Sequence 11, Appl
24	247.5	4.3	328	10	US-09-758-593A-12	Sequence 12, Appl
25	247.5	4.3	329	10	US-09-880-192-62	Sequence 62, Appl
26	247.5	4.3	329	10	US-09-758-593A-1	Sequence 1, Appl
27	243.5	4.2	2444	10	US-09-944-849-2	Sequence 2, Appl
28	243	4.2	752	10	US-09-927-180-2	Sequence 2, Appl
29	240	4.2	671	9	US-09-924-400-306	Sequence 306, App
30	240	4.2	671	9	US-10-012-896-380	Sequence 380, App
31	240	4.2	671	9	US-09-895-793-380	Sequence 380, App
32	240	4.2	671	9	US-09-895-814-380	Sequence 380, App
33	240	4.2	671	10	US-09-825-301-10	Sequence 10, Appl
34	240	4.2	671	10	US-09-759-143-380	Sequence 380, App
35	240	4.2	671	10	US-09-780-669-380	Sequence 380, App
36	240	4.2	671	10	US-09-810-936-306	Sequence 306, App
37	240	4.2	671	10	US-09-822-827-380	Sequence 380, App
38	240	4.2	671	10	US-09-429-755-306	Sequence 306, App
39	237.5	4.1	656	9	US-09-924-400-305	Sequence 305, App
40	237.5	4.1	656	9	US-10-012-896-379	Sequence 379, App
41	237.5	4.1	656	9	US-09-895-793-379	Sequence 379, App
42	237.5	4.1	656	9	US-09-895-814-379	Sequence 379, App
43	237.5	4.1	656	10	US-09-825-301-9	Sequence 9, Appl
44	237.5	4.1	656	10	US-09-759-143-379	Sequence 379, App
45	237.5	4.1	656	10	US-09-780-669-379	Sequence 379, App

ALIGNMENTS

RESULT 1

US-09-509-196A-2

; Sequence 2, Application US/09509196A

; Patent No. US20020037582A1

; GENERAL INFORMATION:

; APPLICANT: DALY, Roger J.

; APPLICANT: SUTHERLAND, Robert L.

; TITLE OF INVENTION: A Potential Effector for the Grb7 Family of Signalling

; TITLE OF INVENTION: Proteins

; FILE REFERENCE: 1871-129

; CURRENT APPLICATION NUMBER: US/09/509,196A

; CURRENT FILING DATE: 2000-03-23

; PRIOR APPLICATION NUMBER: P09388

; PRIOR FILING DATE: 1997-09-23

; PRIOR APPLICATION NUMBER: PCT AU98/00795

; PRIOR FILING DATE: 1998-09-23

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 1074

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-509-196A-2

Query Match	97.1%	Score 5602;	DB 10;	Length 1074;
Best Local Similarity	99.4%	Pred. No. 0;		
Matches 1068;	Conservative 2;	Mismatches 4;	Indels 0;	Gaps 0;
QY	27	IFLHNACSFHAEVNVLLRLRHGADPNARDNNWYTPLEHAAIKGKIDVCIVLLQHGAEPTI	86	
Db	1	IFLHNACSFHAEVNVLLRLRHGADPNARDNNWYTPLEHAAIKGKIDVCIVLLQHGAEPTI	60	
QY	87	RNTDGTALDADPSAKAVLTGEYKKDELLESARGSEKMMALLTPLNVNCHASDGRKS	146	
Db	61	RNTDGTALDADPSAKAVLTGEYKKDELLESARGSEKMMALLTPLNVNCHASDGRKS	120	
QY	147	TPHLAAGYNRVKIVQLLQHGADVHAKDGLVPLHNACSYGHVEYTELIVKKGACVNA	206	
Db	121	TPHLAAGYNRVKIVQLLQHGADVHAKDGLVPLHNACSYGHVEYTELIVKKGACVNA	180	
QY	207	MDLWQFTPLHEAASKNRVEVCSLLSYGADPTLLNCHNKSATDLAPTQLKERLAYEFKG	266	

Db 181 MDLWQFTPLHEAASKNRVEVCSLLSYGADPTLLNCKNKSALDAPTPOLKERLAYEFKG 240
QY HSLLOAAREADVTRIKKHLSLEMVNFKHPOTHTALHCAASPYPKRKQICELLIRKGAN 326
Db 241 HSLLOAAREADVTRIKKHLSLEMVNFKHPOTHTALHCAASPYPKRKQICELLIRKGAN 300
QY INEKTKEFTPLHVASEKANDVVEVVKHAKVNDLNI.GQTSLHRAAYCGHLQTCRLL 386
Db 301 INEKTKEFTPLHVASEKANDVVEVVKHAKVNDLNI.GQTSLHRAAYCGHLQTCRLL 360
QY LSYGCDPNIISLOGFTALONGNENVOQLQEGISLGNSEADROLLLEAAKAGDVETVKKLC 446
Db 361 LSYGCDPNIISLOGFTALONGNENVOQLQEGISLGNSEADROLLLEAAKAGDVETVKKLC 420
QY TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNACSYGHI 506
Db 421 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNACSYGHI 480
QY EVAELLVKGVAVNVADLWKFTPLHEAAAKGYEICKLLQHGADPTKKNRDNTPLDLV 566
Db 481 EVAELLVKGVAVNVADLWKFTPLHEAAAKGYEICKLLQHGADPTKKNRDNTPLDLV 540
QY KDGDTDODLLRGDAALLDAAGKGLARVKLSSPDNVNCRDTQGRHSTPLHLAAGYNIL 626
Db 541 KDGDTDODLLRGDAALLDAAGKGLARVKLSSPDNVNCRDTQGRHSTPLHLAAGYNIL 600
QY EVAEYLLOHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFPLHEA 686
Db 601 EVAEYLLOHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNASLNATDKWAFPLHEA 660
QY AQKGRTOCALHAGADPTLKNQEGOTPLDLSADDDVSALLTAAMPSPALPCYKPOVL 746
Db 661 AQKGRTOCALHAGADPTLKNQEGOTPLDLSADDDVSALLTAAMPSPALPCYKPOVL 720
QY NGVRSFGATADALSSGSPSSLSAASSLDNLSSGSESSVSSVSSGTEGASSLEKKEVP 806
Db 721 NGVRSFGATADALSSGSPSSLSAASSLDNLSSGSESSVSSVSSGTEGASSLEKKEVP 780
QY GVDFTSQFVRNLGLEHLMDFEREQITLDLVEMGHKELKEIGINAYGHRHKLKIGVER 866
Db 781 GVDFTSQFVRNLGLEHLMDFEREQITLDLVEMGHKELKEIGINAYGHRHKLKIGVER 840
QY LLSGQOGLNPYLNTSSGSTILIDLSPDDKEFQSVSEEMQSTVRHRDGGHAGGIFNRY 926
Db 841 LLSGQOGLNPYLNTSSGSTILIDLSPDDKEFQSVSEEMQSTVRHRDGGHAGGIFNRY 900
QY NTLKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAI IHKGFDERHAYIG 986
Db 901 NTLKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAI IHKGFDERHAYIG 960
QY GMFGAGIYFAENSSKNQVYIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLOFSA 1046
Db 961 GMFGAGIYFAENSSKNQVYIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLOFSA 1020
QY MKMAHSPGCHSHVTRGPSVGLALAEYVYIRGEAOAPEYLYIYQIMRPGMVDG 1100
Db 1021 MKMAHSPGCHSHVTRGPSVGLALAEYVYIRGEAOAPEYLYIYQIMRPGMVDG 1074

RESULT 2

US-09-841-835-2

; Sequence 2, Application US/09841835

; Patent No. US20020076795A1

; GENERAL INFORMATION:

; APPLICANT: de Lange, Titia

; APPLICANT: Smith, Susan

; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS

; TITLE OF INVENTION: OF USE THEREOF

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber & Jackson

; STREET: 411 Hackensack Avenue, 4th Floor

; CITY: Hackensack

STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,835
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/196,387
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-230 CIPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1327 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-841-835-2

Query Match

84.8%; Score 4889.5; DB 10; Length 1327;

Best Local Similarity 83.0%; Pred. No. 0;

Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

QY 1 GFGKDVVEYLLQNGASVOARDGGIPLHNACSFHAEVNVNLLRHGADPNARDNNYIT 60
Db 225 GFGKDVVEYLLQNGASVOARDGGIPLHNACSFHAEVNVNLLRHGADPNARDNNYIT 284
QY 61 PLHEAAIKGKIDVCIVLLQHGAEPTTRNTDGRDALDAPSAKAVITGEYKKDELLESA 120
Db 285 PLHEAAIKGKIDVCIVLLQHGAEPTTRNTDGRDALDAPSAKAVITGEYKKDELLESA 344
QY 121 SGNEEKMMALLPLNVNCHASDGRKSTPLHAAAGYNRVIVQLLQHGADVHAKDKGLV 180
Db 345 SGNEEKMMALLPLNVNCHASDGRKSTPLHAAAGYNRVIVQLLQHGADVHAKDKGLV 404
QY 181 PLHNACSYGHEVTELLVKHGACVNAAMLWQFTPLHEAASKNRVEVCSLLSYGADPTLL 240
Db 405 PLHNACSYGHEVTELLVKHGACVNAAMLWQFTPLHEAASKNRVEVCSLLSYGADPTLL 464
QY 241 NCHNSAIDLAPTPOLKERLAYEFKSHLSQAAREADVTRIKKHLSLEMVNFKHPOTHT 300
Db 465 NCHNSAIDLAPTPOLKERLAYEFKSHLSQAAREADVTRIKKHLSLEMVNFKHPOTHT 524
QY 301 ALHCAASPYPKRKQICELLIRKGANINEKTEFTPLHVASEKANDVVEVVKHAKV 360
Db 525 ALHCAASPYPKRKQICELLIRKGANINEKTEFTPLHVASEKANDVVEVVKHAKV 584
QY 361 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDPNIISLOGFTALONGNENVOQLQEGIS 420
Db 585 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDPNIISLOGFTALONGNENVOQLQEGIS 644
QY 421 LGNSEADROLLLEAAKAGDVETVKKLCVTSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
Db 645 IRTSDVDYRLELSEAKAGDLETVKQLCSSQNVNCRDIEGRHSTPLHFAAGYNRVSVVEYLL 704
QY 481 QHGADVHAKDKGLVPLHNACSYGHEVTELLVKHGACVNAAMLWQFTPLHFAAASKNRVE 540
Db 705 QHGADVHAKDKGLVPLHNACSYGHEVTELLVKHGACVNAAMLWQFTPLHFAAASKNRVE 764

```
Qy 541 ICKLLQHGADPTKKNRDGNTPLDLVKDGDITDIDLLRGDAALDAKKGCLARVKKLSS 600
Db 765 ICKLLKKGADPTKKNRDGNTPLDLVKEGDDTDIDLLKGDAAALDAKKGCLARVKKLCT 824
Qy 601 PDNVNCRDQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDGGGLPLHNAASYGHVD 660
Db 825 PENINCRDQGRNSTPLHLAAGYNLEVAEYLLQHGADVNAQDGGGLPLHNAASYGHVD 884
Qy 661 VAAALLIKYNACVNATDKWAFPLHEAAKGRTOLCALLAHGADPTLKNQOGPTDLVS 720
Db 885 IAALLIKYNTCVNATDKWAFPLHEAAKGRTOLCALLAHGADPTMKNQOGPTDLAT 944
Qy 721 ADVSALLTAAMPSPALSCYKPVOLNGVRSPGATADALSSGSPSSLSAASIDNLG 780
Db 945 ADIIRALLIDAMPPEALPTCEKQAT-----VVSASLISPASTPSCLSAASIDNLG 997
Qy 781 SFSESSVSVSGTGEASLSLEKK--EVPGVDFSTQFVRLNGLHLMDFIFEREQITLDVL 838
Db 998 PLAELAVGASNAGDGAAGTERKEGEVAGLDWNIISOFLKSLGLEHLRDIIFETEQITLDVL 1057
Qy 839 VEMGHKEKEIGINAYGHRHKLIGVERLISGOOGLNPYLTLNTSGSCTIILIDSPDKE 898
Db 1058 ADMGHEELKEIGINAYGHRHKLIGVERLISGOOGLNPYLTLNTSGSCTIILIDSPDKE 1117
Qy 899 FOSVEEEMQSTVREHRDGGHAGGIFNRYNLIKIQVCNKKLWERYTHRRKEVSEENHHA 958
Db 1118 YQSVEEEMQSTIREHRDGGNAGGIFNRYNIRIQVNVNKKLRERFCHQKEVSEENHHA 1177
Qy 959 NERMLFPHGSPFNALIHKGFDERRHAYIGMGFAGIYFAENSCKSNQYVYGGGTGCPVH 1018
Db 1178 NERMLFPHGSPFNALIHKGFDERRHAYIGMGFAGIYFAENSCKSNQYVYGGGTGCPVH 1237
Qy 1019 KORSVCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSTVGRPSVNGLALAEVYVIRG 1078
Db 1238 KORSVCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSTVGRPSVNGLALAEVYVIRG 1297
Qy 1079 EQAYPEYLITYQIMRPE 1095
Db 1298 EQAYPEYLITYQIMRPE 1314

RESULT 3
US-09-964-899-41
; Sequence 41, Application US/09964899
; Patent No. US2002017446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; FILE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-41

Query Match 62.3%; Score 3596; DB 9; Length 802;
Best Local Similarity 65.1%; Pred. No. 2.3e-253;
Matches 720; Conservative 5; Mismatches 5; Indels 376; Gaps 5;

Qy 1 GFGRKDVVEYLLQNGASVARDGGGLPLHNAACSFHAEVNNLLLRHGADPNARDNNWT 60
Db 67 GFGRKDVVEYLLQNGAVQARDGGGLPLHNAACSFHAEVNNLLLRHGADPNARDNNWT 126
Qy 61 PLHEAAIKGIDVCIVLLQHGAEPTIRNTDGRALDADPSAKAVLTGE-----YKKDEL 115
```

```
Db 127 PLHEAAIKGIDVCIVLLQHGAEPTIRNTDGRALDADPSAKAVLTGKSVSGYSRKPV 186
Qy 116 LES-ARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK 174
Db 187 KNNLARSGNEEKMMALLTPLNVNCHASDGR----- 216
Qy 175 DKGDVPLHNACSYGHEVTELLVKGACVYNAMDLQWFTPLHEAASKNRVYEVCSLLLSYG 234
Db 217 -----KHCACVYNAMDLQWFTPLHEAASKNRVYEVCSLLLSYG 252
Qy 235 ADPTLLMCHNKSALDAPTQPKERLAYEPKGHSLLOAAREADVTRIKKHLSELMVNFKH 294
Db 253 ADPTLLMCHNKSALDAPTQPKERLAYEPKGHSLLOAAREADVTRIKKHLSELMVNFKH 312
Qy 295 POTHETALHCAASAPYPKQKQICELLLRKCANINEKTEFLTLPHVASEKAHNVVEVV 354
Db 313 POTHETAL----- 320
Qy 355 KHEAKVNALDNLGOTSLHRAAYCGHLQTCRLLLSYGDPMNIIISLOGFTALQMGNNVQOL 414
Db 321 ----- 320
Qy 415 LOEGISLGNSEADROLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRV 474
Db 321 -----KLCTVQSVNCRDIEGRQSTPLHFAAGYNRV 351
Qy 475 VVEYLLQHGADVHAKDGGGLVPLHNACSYGHEVAELLVKHGAVNVNADLWKTFTPLHEAA 534
Db 352 VVEYLLQHGADVHAKDGGGLVPLHNACSYGHEVAELLVKHGAVNVNADLWKTFTPLHEAA 411
Qy 535 AKGYEICKLLQHGADPTKKNRDGNTPLDLVKDGDITDIDLLRGDAALDAKKGCLAR 594
Db 412 AKGYEICKLLQHGADPTKKNRDGNTPLDLVKDGDITDIDLLRGDAALDAKKGCLAR 471
Qy 595 VKKLSSPDNVNCRDQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDGGGLPLHNA 654
Db 472 VKKLSSPDNVNCRDQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDGGGLPLHNA 531
Qy 655 SYGHVDVAAALLIKYNACVNATDKWAFPLHEAAOKGRTOLCALLAHGADPTLKNQOG 714
Db 532 SYG----- 534
Qy 715 PLDLVSADVSALLTAAMPSPALSCYKPVOLNGVRSPGATADALSSGSPSSLSAAS 774
Db 535 ----- 534
Qy 775 LDNLGSESELSSVSVSGTGEASLSLEKKYVPYDFSTQFVRLNGLHLMDFIEREQIT 834
Db 535 -----IT 536
Qy 835 LDVLVEMGHKELKEIGINAYGHRHKLIGVERLISGOOGLNPYLTLNTSGSCTIILDSP 894
Db 537 LDVLVEMGHKELKEIGINAYGHRHKLIGVERLISGOOGLNPYLTLNTSGSCTIILDSP 596
Qy 895 DDKEFQSVVEEEMQSTVREHRDGGHAGGIFNRYNLIKIQVCNKKLWERYTHRRKEVSEEN 954
Db 597 DDKEFQSVVEEEMQSTVREHRDGGHAGGIFNRYNLIKIQVCNKKLWERYTHRRKEVSEEN 656
Qy 955 HNHNARMFLPHGSPFNALIHKGFDERRHAYIGMGFAGIYFAENSCKSNQYVYGGGTG 1014
Db 657 HNHNARMFLPHGSPFNALIHKGFDERRHAYIGMGFAGIYFAENSCKSNQYVYGGGTG 716
Qy 1015 CPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSTVGRPSVNGLALAEV 1074
Db 717 CPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSTVGRPSVNGLALAEV 776
Qy 1075 IYRGEQAYPEYLITYQIMRPEGMVDG 1100
Db 777 IYRGEQAYPEYLITYQIMRPEGMVDG 802

RESULT 4
```

US-09-841-835-10
; Sequence 10, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 949 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-841-835-10

Query Match 53.6%; Score 3093; DB 10; Length 949;
Best Local Similarity 81.5%; Pred. No. 1.1e-216;
Matches 585; Conservative 60; Mismatches 71; Indels 2; Gaps 1;

QY 1 GFGKDVVEYLLQNGASVOARDGGGLPLHNACSFGEAEVWVLLLRHGDADPNARDNNYNT 60
DB 225 GFGKDVVEYLLQNGANVANRDGGGLPLHNACSFGEAEVWVLLLRHGDADPNARDNNYNT 284
QY 61 PLHEAAIKGIDVCIVLLQHGAEPTIRNTDGRALDADPSAKAVLTGEYKKDELLESAR 120
DB 285 PLHEAAIKGIDVCIVLLQHGADPNRNTDGSALDADPSAKAVLTGEYKKDELLEAAR 344
QY 121 SGNEEKWALLPLPNVCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLV 180
DB 345 SGNEEKWALLPLPNVCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLV 404
QY 181 PLHNACSYGHEVTELLVKGACVAMDLWOFTPLHEAAKRNVEVCSSLLLSYGADPTLL 240
DB 405 PLHNACSYGHEVTELLVKGACVAMDLWOFTPLHEAAKRNVEVCSSLLLSYGADPTLL 464
QY 241 NCHNSAIDLAPTPOLKERLAYEFKSHLSQAAREADVTRIKKHLSEMVNFKHPOTHE 300
DB 465 NCHGKSAVDMAPTPELRRLTYEFKSHLSQAAREADVTRIKKHLSEMVNFKHPOTHE 524
QY 301 ALHCAAAASPYPRKQICELLRLKGNINPKTEFTPLHVASEKAHNDVVEVVKHEAV 360
DB 525 ALHCAVASLHPRKQVTELLRLKGNANVEKNKDFMTPLHVAAERAHNDVVEVVKHEAV 584
QY 361 NALDNLGOTSLHRAAYCGHLQTCRLLSYGCDPNIISLQGTALQMGNEVQVLLQEGIS 420

DB 585 NALDTIGOTALHRAALAGHLQTCRLLSYGSDPSIISLQGTAAQMGNEAVQVLLSESTP 644
QY 421 LGNSEADRLLEAAKAGDVETVKKLCTVQSYNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
DB 645 IRTSDVYRLLEASKAGDLETVKOLCSSQNVNCRDIEGRHSTPLHFAAGYNRVSVVEYLL 704
QY 481 QHGADVHAKDKGGLVPLHNACSYGHEVTELLVKGAVVNVADLWKETPLHEAAAKGKYE 540
DB 705 HHGADVHAKDKGGLVPLHNACSYGHEVTELLVKGAVVNVADLWKETPLHEAAAKGKYE 764
QY 541 ICKLLQHGADPTKKNRDGNTPLDVRDGDPTDIDLLRGGDAALDDAAKKGCLARVKLLSS 600
DB 765 ICKLLQHGADPTKKNRDGNTPLDVRDGDPTDIDLLRGGDAALDDAAKKGCLARVKLLSS 824
QY 601 PDVNCRTQGRHSTPLHLAAGYNLNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH-- 658
DB 825 PENINCRDTQGRNSTPLHLAAGYNLNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGCL 884
QY 659 VDVAALLIKYNACVNTDKWAFPLHEAAOKGRQTOLCALLAHGADPTLNQEGQTPL 716
DB 885 ARVOKLCTPENINCRTQGRNSTPLHLAAGYNLNLEVAEYLLQHGADVNAQDKGGLIPL 942

RESULT 5

US-09-841-835-8
; Sequence 8, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 673 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-841-835-8

Query Match 33.3%; Score 1918.5; DB 10; Length 673;
Best Local Similarity 80.9%; Pred. No. 1.7e-131;
Matches 364; Conservative 42; Mismatches 43; Indels 1; Gaps 1;

```
QY 1 GFGKDVVEYLLONGASVOARDGGLIPLHNACSFGEAEVNVNLLRHGADPNARDNNYT 60
Db 225 GFGKDVVEYLLONGASVHARDGGLIPLHNACSFGEAEVNVNLLRHGADPNARDNNYT 284
QY 61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRALDADPSAKAVLTGEYKKDELLEAAR 120
Db 285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGRSALDADPSAKAVLTGEYKKDELLEAAR 344
QY 121 SNEEKMMALLTPLNVNCHASGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDGLV 180
Db 345 SNEEKMMALLTPLNVNCHASGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDGLV 404
QY 181 PLHNACSYGHEVTELLYKHGACVAMDLWOFTPLHEAASKNRVEVCISLLSYGADPTLL 240
Db 405 PLHNACSYGHEVTELLYKHGACVAMDLWOFTPLHEAASKNRVEVCISLLSHGADPTLV 464
QY 241 NCHNKSADLAPTQPKERLAYEFKSHLSLQAAAREADYTRIKKHLSELMVNFKHPQTHET 300
Db 465 NCHGKSADMAPTPELRLERTYEFKSHLSLQAAAREADLAKVKTALLETINFKQPOSHET 524
QY 301 ALHCAASAPYPRKQICELLRLKGANINEKTEFLTPHVAASEKAHNDVVEVVKHEAKV 360
Db 525 ALHCAVASLHPKQVTELLRLKGANINEKNDFTPLHVAERAHNDVMEVLHKGAKM 584
QY 361 NALDNLGOTSLHRAAYCGHLOTCTRLLSYGCDPNIISLOGFTALQMGNEVQQLQEGIS 420
Db 585 NALDTLCQTHALHRAALAGHLOTCTRLLSYGSDDPSIISLOGFTAAQMGNEAVQQLLSVYG 644
QY 421 LGNSEADROLLEAAKAGDVETVKKLTCTVOS 450
Db 645 SDPSIISLOGFTAAQMGNEAVQQLLSGHS 673

RESULT 6
US-09-964-899-43
; Sequence 43, Application US/09964899
; Patent No. US2002017446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 1724
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-43

Query Match 13.8%; Score 798.5; DB 9; Length 1724;
Best Local Similarity 25.2%; Pred. No. 1.9e-49;
Matches 279; Conservative 167; Mismatches 380; Indels 279; Gaps 28;

QY 3 GRKDVVEYLLONGASVOARDGGLIPLHNACSFGEAEVNVNLLRHGADPNARDNNYTPL 62
Db 47 GHVEVSWELLQREANVDAATKKGNTALHIAISLAGAEVVKVLTNGANVNAQSQNGFTPL 106
QY 63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRALDADPSAKAVLTGEYKKDELLEAARS 122
Db 107 YMAAQENHLEVYKFLDNGASQSLSATGDTPLAVA-----LQQG 146
QY 123 NEEKMMALLTPLNVNCHASGRKSTP-LHLAAGYNRVKIVOLLQHGADVHAKDGLV 181
Db 147 HDQVVSLLLE-----NDTKGVKPLPALHTAARKDDYTKAALLQNDNNADVESKSGFT 200
QY 182 LHNACSYGHEVTELLYKHGACVAMDLWOFTPLHEAASKNRVEVCISLLSYGADPTLLN 241
```

```
Db 201 LHIAAHYGNINVAITLLNRAAAVDFTARNDDITPLHVASKRGANMYKLLDRGA----- 254
QY 242 CHNKSATDLAPTQPKERLAYEFKSHLSLQAAAREADYTRIKKHLSELMVNFKHPQTHETA 301
Db 255 -----KIDAKTRDGL-----TP 266
QY 302 LHCAASAPYPRKQICELLRLKGANINEKTEFLTPHVAASEKAHNDVVEVVKHEAKVN 361
Db 267 LHCGARSGH---EQVEMLLDRAAPILSKTKNGLSPLHMAATQGDHLNCVQLLQHNVPVD 323
QY 362 ALDNLGOTSLHRAAYCGHLOTCTRLLSYGCDPNIISLOGFTALQMGNEVQQLQEGISL 421
Db 324 DVTNDYLTALHVAACHGHYKAVKVLDDKAPNAKALNGFTPLHI----- 368
QY 422 GNSEADROLLEAAKAGDVETVKKLTCTVOSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLQ 481
Db 369 -----ACKNRKIK-VMEILLKHGASIQAVTERGETALHMAASQGAQEVRYLVQ 416
QY 482 HGADVHAKDGLVPLHNACSYGHEVVAELLYKHGAVNVNADLWKFTPLHEAAKGYEI 541
Db 417 DGAQVEAKAKDDQTPHLISARLGRKADIVQQLLOQOGASPAATTSYGTPLHLSAREGHEDV 476
QY 542 CKLLLOHGCADPTKKNRGNTPLDLVKDGDTDIQLLRGDAALLD-----AAKKGCLAR 594
Db 477 AAFLLDHGASLSITTKSGLTPLHVAAHYDNOKVALL-----LDDQASPHAAKNG---- 527
QY 595 VKLSSPDNVNCRDTQGRHSTPLHAAAGYNLNVAEYLLQHGADVNAQDKGLIPLHNAA 654
Db 528 -----YTPHLIAAKKNQMDIATTLLEYGADANAVTROGIAVHLAA 568
QY 655 SYGHVDVAALLIKYNACVNATDKWAFTPLEAOKGRQTQLCALLANGADPTLNQ----- 710
Db 569 QEGHVDVMSLLGRNANVNLNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVDAQTQKVGYP 628
QY 711 ---EGOTPLDLSADDVSALLTAAMPSPALPCYKQVNLGVRSFGATADALSSGSSPS 767
Db 629 PLPHGKECVHLVTANGYTPHQAQQGHT-----HIINVLQNNASPNELTVTVTEKH 681
QY 768 SLASAASS-----LDNLSSGSFSELSSVSSGTEGASSLEK-----KEYPGV-DFSITQFV 816
Db 682 KMNVPETWNEVLNDSDDVEVRKANAPEMLSDEGYISDVEEGNRCWTYKIPKVOEFTVK--- 738
QY 817 RNLGLEHMLDIPEREQITDLVIVE-----MGHKLKEIGINAY----- 854
Db 739 -----TDTFKREAFDVGLISTSGEDAMTGDYKLGPDQLKELGDDSLPABGYMGF 790
QY 855 -----GHRHKLKQVE-----RLISQQQGLNPYL 878
Db 791 SLGARSARFLVSFMDVARGGSMRGSRHHGMRILIPPRCKTAPTRITCRVLRKHLANPPP 850
QY 879 TLNTSGSGTILHDLSPDQKEFOS-VEEEMQSTVREHROGHAGGFNRYNLIKQKCNK 937
Db 851 MVEGEGLASRLVEMGMPAGAPLGPVIVEIP-----HFGSMRGKRERELIVLRSNG 900
QY 938 KLWE--RYTHRRKEVSE-----ENHNHANERMLFHCSFPVNAIHKGFDERHAYIGGFG 990
Db 901 ETWKEHQFDSKNEDUTELLNGDELDLSPEEL---GKKRICRILITKDPQ----- 947
QY 991 AGTYFAENS---SKSNQVYVIGGG 1012
Db 948 ---YFAVWSRIQESNQ---IGPEGG 967

RESULT 7
US-09-908-711-70
; Sequence 70, Application US/09908711
; Patent No. US20020045230A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA128
; CURRENT APPLICATION NUMBER: US/09/908,711
```

```

; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (148)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (167)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (169)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (258)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (396)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (413)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (414)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (417)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (418)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (421)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-908-711-70

```

```

Query Match      6.8%; Score 395; DB 10; Length 426;
Best Local Similarity 27.9%; Pred. No. 5e-21;
Matches 123; Conservative 52; Mismatches 160; Indels 106; Gaps
7;

268 SLLOAAR----EADVTRIKKHSLEMVNFKHPQTHETALHCAASPPPKRKQICELLRK 323
||| ||| | : : : | | ||| | : : |||
44 SYLRARAGHLEKALDYIKNGVDINICN---QNGNALHLASKEGH--VEVVSSELLQR 96
||| ||| | : : : | | ||| | : : |||
324 GANINIEKTEFLTPHVASEKAHNDVVVVVKEAKVNALNDNLGQTSLHRAAYCGHLQTC 383
||| ||| | : : : | | ||| | : : |||
97 EANVDAATKKGNTALHIASIAGQAEVVVVLVTGANCNVNAQSQNGFTPLYMAXENHLEVV 156
||| ||| | : : : | | ||| | : : |||
384 RLLLSYCGDPNIISLQGFT----ALOMGNENVQOOLQEGISLGNSEADROLLEAAKAGDV 439
||| ||| | : : : | | ||| | : : |||
157 KFLLDNGASQXLXTBDCFTPLAVALQOQHDQVVVSLLE----- 194
||| ||| | : : : | | ||| | : : |||
440 ETVKKLCITVQSVNCRDIEGRQSTP-LHFAAGYNRVSVVEYLLQHGADVHAHKDGLVPLH 498
||| ||| | : : : | | ||| | : : |||
195 -----NDTKGKVRPLPAHIAARKDDTKAAALLQNDNNADVESKSGTPLH 240
||| ||| | : : : | | ||| | : : |||
499 NACSIGHYEVABELLVKGAVVNVADLWKFPLHEAAKAKYEICKLLLOHGADPTKKNRD 558
||| ||| | : : : | | ||| | : : |||
241 IAAHYGNINVATLLNLRXAAVDFTARNDITPLHVASKRGNNANMVKLLLDREGAKIDAKTRD 300
||| ||| | : : : | | ||| | : : |||
559 GNTPLDLVKDGDTDQDILLRGDAALLDAAKGCLARVKKLSLSPDNVNCRDTOGRHSTPLH 618
||| ||| | : : : | | ||| | : : |||
301 G-----LTPLH 306
||| ||| | : : : | | ||| | : : |||
619 LAAGYNNLEVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNATDKW 678
||| ||| | : : : | | ||| | : : |||
307 CGARSGHEQVVEMLLDRAAPILLSKTKNGLSPUHMATQGDHLNCVQLLLQHNVPVDVNTD 366
||| ||| | : : : | | ||| | : : |||
679 AFTPLHAAQKGRGTQLCALLL 699
||| ||| | : : : | | ||| | : : |||

```


;; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
;; FILE REFERENCE: 210121.534C2
;; CURRENT APPLICATION NUMBER: US/09/895,793
;; CURRENT FILING DATE: 2001-06-29
;; NUMBER OF SEQ ID NOS: 982
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 378
;; LENGTH: 1719
;; TYPE: PRT
;; ORGANISM: Homo sapien
US-09-895-793-378

Query Match 6.6%; Score 379.5; DB 9; Length 1719;
Best Local Similarity 20.9%; Pred. No. 5.6e-19;
Matches 205; Conservative 106; Mismatches 335; Indels 337; Gaps 23;
QY 143 GRKSTPLHLAGYNNRYKIVQLLQ-HGADVHAKDGLVPLHNACSYGHYEVTELLVKG 201
DB 137 GEDLDKLRHAAWGWKVPKRDILVMLRDTVNNKKOKRKTALHLASANGSEVVKLLDR 196
QY 202 ACVNAMDLWQFTPLHBAASKNRVEVCSLLSYGADPTLLNCHNKSAIDLAPTOLKERLA 261
DB 197 COLNVLDNKKRTALIKAVQCEDECALMLEHGTDPNI----- 234
QY 262 YEFKGSLLQAAREADVTIRIKKHSLEWVNFKHPQTH-ETALHCAASPYPKRKQICELL 320
DB 235 -----PDEYGNITLHYAI-----YNEDKLMKAL 258
QY 321 LRKGANINEKTEFTPLHVAASEKAHNDVVEVVKHEAKVNALDNLGOTSLHRAAYCGHL 380
DB 259 LLYGADIESKNKHGLTPLLGVHEQKQVVKFLIKKANLNALDRYGRITALILAVCCGSA 318
QY 381 QTCRLLLSYGDNPNIISIQGFTALQMG-----NENVOQLQOE-----GISLGNSEAD-- 427
DB 319 STVSLLEQINIDVSSQDLSGQTAREYAVSSHVVICOLLSDYREKOMLKISSENSPENV 378
QY 428 -----ROLLEAAKAGDVETVKKLCTVQSVN-----CRD----- 455
DB 379 SRTNRPRTTHVVEVDSMPAASVVKPFGLRSMKMGKWCCHCPCRESGKSNVGTSGDHD 438
QY 456 ----- 455
DB 439 DSAMKTLRSKMGKWCCHCPCRGSGKSNVGSAGSDHDSAMKTLRNKMGKWCCHCPCCR 498
QY 456 -----IEGRQSTPLHFAAGYNNRYKIVQLLQ-HGADVHAKD 490
DB 499 GSKSKVGAMGDYDDSAFMERPHVHVGEDLDKLRHAAWGWKVPKRDILVMLRDTVNNKKD 558
QY 491 KGLVPLHNACSYGHYEVTELLVKGAVNVADLWKFPLHBAASKNRVEVCSLLSYGADPTLLNCHNKSAIDLAPTOLKERLA 550
DB 559 KOKRTALHLASANGSEVVKLLDRRCQLNVLDNKKRTALIKAVQCEDECALMLEHGT 618
QY 551 DPTKKNRDNTPDLVKGDTDIQDILLRGDAALLDAAKGLARVKKLSPPDNVNCRDITQ 610
DB 619 DNPIDPEYGN----- 628
QY 611 GRHSTPLHLAGYNNRYKIVQLLQ-HGADVHAKDGLVPLHNACSYGHYEVTELLVKG 670
DB 629 -----TTLHYAIYNEDKLMKALLLYGADIESKNKHGLTPLLGVHEQKQVVKFLIKKKA 684
QY 671 CVNATDKWAFTPLHBAASKNRVEVCSLLSYGADPTLLNCHNKSAIDLAPTOLKERLA 730
DB 685 NLNALDRYGRITALILAVCCGSAIYVSLLEQINIDVSSQDLSGQTAREYAVSSHVV----- 741
QY 731 AMPPSALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASSDNLNLSGFSSELSSVVS 790
DB 742 -----COLLSYDEKQKMLKISSE-----NSNPEQDLKLTSEESQRFKGS----- 781
QY 791 SSGTEGASLEKKEKVPVDFSTQFVRNIGLEHMLDIFEREQITLDLVYEMGHKELKEIG 850
DB 782 -----ENSOPEKMSQPEINKD-----GDRE-----VE-----EEMKHE 811
QY 851 INAYGHRHKLIGVERLISGQOGLNPYLTNTSGSTILIDLSPDDKEFQSEVEEMQSTV 910

DB 812 SNNVGLLENLTNGV-TAGNGDNGLIPOQRKSR-----PENQOFPDNESEYHRI 859
QY 911 REHRDGGHAGGIFNRYNLIKIKQKCNK-----KLWERYTHRRKKEVSEENHH----- 957
DB 860 CE-----LVSDYKEKQMPKYSSENSNPQDLKLTSEESQRLGSENGOPELEPNM 910
QY 958 ANERMLFHGSPFVNALIIHKGEDE 980
DB 911 AIEEMKKHGS-----THVGFPE 927
RESULT 10
US-09-895-814-378
; Sequence 378, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 378
; LENGTH: 1719
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-895-814-378

Query Match 6.6%; Score 379.5; DB 9; Length 1719;
Best Local Similarity 20.9%; Pred. No. 5.6e-19;
Matches 205; Conservative 106; Mismatches 335; Indels 337; Gaps 23;

QY 143 GRKSTPLHLAGYNNRYKIVQLLQ-HGADVHAKDGLVPLHNACSYGHYEVTELLVKG 201
DB 137 GEDLDKLRHAAWGWKVPKRDILVMLRDTVNNKKOKRKTALHLASANGSEVVKLLDR 196
QY 202 ACVNAMDLWQFTPLHBAASKNRVEVCSLLSYGADPTLLNCHNKSAIDLAPTOLKERLA 261
DB 197 COLNVLDNKKRTALIKAVQCEDECALMLEHGTDPNI----- 234
QY 262 YEFKGSLLQAAREADVTIRIKKHSLEWVNFKHPQTH-ETALHCAASPYPKRKQICELL 320
DB 235 -----PDEYGNITLHYAI-----YNEDKLMKAL 258
QY 321 LRKGANINEKTEFTPLHVAASEKAHNDVVEVVKHEAKVNALDNLGOTSLHRAAYCGHL 380
DB 259 LLYGADIESKNKHGLTPLLGVHEQKQVVKFLIKKANLNALDRYGRITALILAVCCGSA 318
QY 381 QTCRLLLSYGDNPNIISIQGFTALQMG-----NENVOQLQOE-----GISLGNSEAD-- 427

wed Feb 12 16:08:51 2003

```

; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 378
; LENGTH: 1719
; TYPE: PRT
; ORGANISM: Homo sapien
; US-0759-143-378

Query Match 6.6%; Score 379.5; DB 10; Length 1719;
Best Local Similarity 20.9%; Pred. No. 5.6e-19;
Matches 205; Conservative 106; Mismatches 335; Indels 337; Gaps 23;

QY 143 GRKSTPLHLAAGYNNRVRKIVQLLIQ-HGADVHAKDGLVPLHNAACSYGHYEVTELLVKHG 201
DB 137 GEDLDKHLRAAWGKVPKDLIVMLRDTDVNKKDKQKRTALHLASANGSEVVKLLDDR 196
QY 202 ACYNMADLWFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSADLAPTQLKERLA 261
DB 197 CQLNVLDNKKRTALIKAVQCOEDECALMLLEHGDPI 234
QY 262 YEFKHSILQAAAREADVTRIKKHLSEWVNFKHPQTH-ETALHCAAASPYPRKQICE 320
DB 235 -----PDEYGNITLHVAI---YNEDKLMAKAL 258
QY 321 LRKGANINEKTEFLTPLHVAASEKAHNDVVEVVKHAKVNALDNLGOTSLSHRAAYCCHL 380
DB 259 LLYGADTESKNKHGLTPLLGVHEQKQVVKFLIKKANLALDRYGRRTALILAVCCGSA 318
QY 381 QTCRLLISYCGDPNLIISLOGETALQMG---NENWQQLQOR-----GISLSEAD-- 427
DB 319 SIVSLLEONTIDVSSQDLSGOTAREYAVSSHVHHVICQLLSYKEKQMLKISSENSPENV 378
QY 428 -----ROLLEAAKAGDYETVKKLTQVOSN-----CRD----- 455
DB 379 SRTRNKRPTHMVVEVDSMPAASSVVKPGLRSKMGKWCRCFPCCRESGKSNVGTSGDHD 438
QY 456 ----- 455
DB 439 DSAMKTLRSKMGKWCRCFPCCRSKGNVSGASGDHDSAMKTLRNKMGKWCCHCFPCR 498
QY 456 -----IEGRQSTPLHFAAGYNRVSVVEYLLQ-HGADVHAKD 490
DB 499 GSGKSKVGAWGDYDSDAFMEPRYHVRGEDLDKHLRAAWGKVPKDLIVMLRDTDVNKKD 558
QY 491 KGGVPLHNAACSYGHYEVAEALLVKHGVVNVADLWKFTPLHEAAAKGYEICKLLLOHGA 550
DB 559 KOKRTALHLASANGSEVVKLLDRRCQLNVLDNKKRTALIKAVQCOEDECALMLLEHGT 618
QY 551 DPTKKNRGNTPLDLVKGDDTDIQLLRGDAALLDAAKKGLARVKKLSPPDNVNCRTQ 610
DB 619 DPNIPDEYGN----- 628
QY 611 GRHSTPLHLAAGYNNRVRKIVQLLIQ-HGADVHAKDGLVPLHNAACSYGHVDAALLIKYNA 670
DB 629 ----TTLHVAIYNEDKLMKALLLYGADIESKNKHGLTPLLGVHEQKQVVKFLIKKKA 684
QY 671 CVNATDKWAFTPLHEAAKGRQTOLCALLAHGADPTLNQEGOTPLDLVSDADDVSALLTA 730
DB 685 NLNALDRYGRRTALILAVCCGSAIVSLLEQNIIDVSSQDLSGOTAREYAVSSHVHI--- 741
QY 731 AMPPSALPCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLDNLGSGFSELSSVVS 790
DB 742 ----COLLSYKEKQMLKISSE-----NSNPEODLKITSEESORFKGS----- 781
QY 791 SSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIG 850
DB 782 ----ENSQPEKMSQEPENKND-----GDRE-----VE---EEMKKHE 811
QY 851 INAYGRHRLIKVERLISGQOGLNPLYLNTSGSTILIDSPDDKEFQSYEEEMQSTV 910
DB 812 SNNVGLLENLTNGV-TAGNDNGNLIPORKSRT-----PENQOPPDNESEYHRI 859
QY 911 REHRDGGHAGGIENRYNLIKIOKVCNK-----KLWERYTHRRKEVSEENHHN----- 957
DB 860 CE-----LVSDYKEKQMPKYSSENSNPEODLKITSEESORLEGSENGQPELENFM 910
QY 958 ANERMLFHGSPFVNAIHKGFDE 980
DB 911 AIEEMKKHGS-----THVGFE 927

RESULT 11
US-0759-143-378
; Sequence 378, Application US/09759143
; Patent No. US2002022248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.

```

```

; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 378
; LENGTH: 1719
; TYPE: PRT
; ORGANISM: Homo sapien
; US-0759-143-378

Query Match 6.6%; Score 379.5; DB 10; Length 1719;
Best Local Similarity 20.9%; Pred. No. 5.6e-19;
Matches 205; Conservative 106; Mismatches 335; Indels 337; Gaps 23;

QY 143 GRKSTPLHLAAGYNNRVRKIVQLLIQ-HGADVHAKDGLVPLHNAACSYGHYEVTELLVKHG 201
DB 137 GEDLDKHLRAAWGKVPKDLIVMLRDTDVNKKDKQKRTALHLASANGSEVVKLLDDR 196
QY 202 ACYNMADLWFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSADLAPTQLKERLA 261
DB 197 CQLNVLDNKKRTALIKAVQCOEDECALMLLEHGDPI 234
QY 262 YEFKHSILQAAAREADVTRIKKHLSEWVNFKHPQTH-ETALHCAAASPYPRKQICE 320
DB 235 -----PDEYGNITLHVAI---YNEDKLMAKAL 258
QY 321 LRKGANINEKTEFLTPLHVAASEKAHNDVVEVVKHAKVNALDNLGOTSLSHRAAYCCHL 380
DB 259 LLYGADTESKNKHGLTPLLGVHEQKQVVKFLIKKANLALDRYGRRTALILAVCCGSA 318
QY 381 QTCRLLISYCGDPNLIISLOGETALQMG---NENWQQLQOR-----GISLSEAD-- 427
DB 319 SIVSLLEONTIDVSSQDLSGOTAREYAVSSHVHHVICQLLSYKEKQMLKISSENSPENV 378
QY 428 -----ROLLEAAKAGDYETVKKLTQVOSN-----CRD----- 455
DB 379 SRTRNKRPTHMVVEVDSMPAASSVVKPGLRSKMGKWCRCFPCCRESGKSNVGTSGDHD 438
QY 456 ----- 455
DB 439 DSAMKTLRSKMGKWCRCFPCCRSKGNVSGASGDHDSAMKTLRNKMGKWCCHCFPCR 498
QY 456 -----IEGRQSTPLHFAAGYNRVSVVEYLLQ-HGADVHAKD 490
DB 499 GSGKSKVGAWGDYDSDAFMEPRYHVRGEDLDKHLRAAWGKVPKDLIVMLRDTDVNKKD 558
QY 491 KGGVPLHNAACSYGHYEVAEALLVKHGVVNVADLWKFTPLHEAAAKGYEICKLLLOHGA 550
DB 559 KOKRTALHLASANGSEVVKLLDRRCQLNVLDNKKRTALIKAVQCOEDECALMLLEHGT 618
QY 551 DPTKKNRGNTPLDLVKGDDTDIQLLRGDAALLDAAKKGLARVKKLSPPDNVNCRTQ 610
DB 619 DPNIPDEYGN----- 628
QY 611 GRHSTPLHLAAGYNNRVRKIVQLLIQ-HGADVHAKDGLVPLHNAACSYGHVDAALLIKYNA 670
DB 629 ----TTLHVAIYNEDKLMKALLLYGADIESKNKHGLTPLLGVHEQKQVVKFLIKKKA 684
QY 671 CVNATDKWAFTPLHEAAKGRQTOLCALLAHGADPTLNQEGOTPLDLVSDADDVSALLTA 730
DB 685 NLNALDRYGRRTALILAVCCGSAIVSLLEQNIIDVSSQDLSGOTAREYAVSSHVHI--- 741
QY 731 AMPPSALPCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLDNLGSGFSELSSVVS 790
DB 742 ----COLLSYKEKQMLKISSE-----NSNPEODLKITSEESORFKGS----- 781
QY 791 SSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIG 850
DB 782 ----ENSQPEKMSQEPENKND-----GDRE-----VE---EEMKKHE 811

```

QY 851 INAYGHRHLKIGVERLISGOQGLNPLYTLNTSGSTILIDLSPDDKEFQVEEEMOSTV 910
Db 812 SNNVGLLENITNGV-TAGNGDGLIPQKRSRT-----PENQOFPDNESEYHRI 859
QY 911 REHRDGGHAGGIFNRYNLIKTKQVCK-----KLWERYTHRRKEYSEENHNH----- 957
Db 860 CE-----LVSDYKEKMPKYSSENSPEQDLKLTSEESQRLGSENGQPELENFM 910
QY 958 ANERMLFHGSPFVNAIHKGFDE 980
Db 911 AIEEMKKHGS-----THVGPE 927
RESULT 12
US-09-780-669-378
; Sequence 378, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 378
; LENGTH: 1719
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-780-669-378

Query Match 6.6%; Score 379.5; DB 10; Length 1719;
Best Local Similarity 20.9%; Pred. No. 5.6e-19;
Matches 205; Conservative 106; Mismatches 335; Indels 337; Gaps 23;
QY 143 GRKSTPLHLAAGNRYKIVQLLLO-HGADVHAKDGLVPLHNACSYGHYEVTELLVKHG 201
Db 137 GEDLDKLHRAAMWGKVPKRDILVMLRDTDVNKKDKQKRTALHLASANGNSEVVKLLDRR 196
QY 202 ACVNMDLMOFTPLHFAAGNRYKIVQLLLO-HGADVHAKDGLVPLHNACSYGHYEVTELLVKHG 201
Db 197 COLNVLDNKKRTALIKAVOCQDECALMLEHGTDPNI----- 234
QY 262 YEFKSHLLOAAREADVTRIKKHSLEMVNFVKHPOTH-ETALHCAAAAPYPRKQICELL 320
Db 235 -----PDEYGNITLHVAI-----YNEKLMKAKAL 258
QY 321 LRKGANINEKTEFTPLHFAAGNRYKIVQLLLO-HGADVHAKDGLVPLHNACSYGHYEVTELLVKHG 380
Db 259 LLYGADIESKNKHLGTPLLGVHGEQKQVVKPLIKKANLALDRYGRITALILAVCCGSA 318
QY 381 QTCRLLSYGCDPNIISLGFTALQMG-----NENVQQLQOE-----GISLGNSEAD-- 427

Db 319 SIIVSLLLEQNDIVSSODLSQOTAREYAVSSHVHVICQLLSDYKEKQMLKISSSENSPENV 378
QY 428 -----POLLEAAKAGDVETVKKICTVOSVN-----CRD----- 455
Db 379 SRTNKPRTMVEVDMPAASVKKPFGRLSRKMGKWCRCPCPCRESGKSNVGTSGDHD 438
QY 456 ----- 455
Db 439 DSAMKTLRSKMGKWCRCPCPCRGSGKSNVGASGDHDDSAMKTLRNMKGKWCCHCFPCR 498
QY 456 -----TEGROSTPLHFAAGNRYKIVQLLLO-HGADVHAKD 490
Db 499 GSGKSVGAWGDVDDSAFMEPRYHVRGEDLDKLHRAAMWGKVPKRDILVMLRDTDVNKKD 558
QY 491 KGLVPLHNACSYGHYEVAELLVKHGAVNVADLMKFTPLHFAAGNRYKIVQLLLO-HGADVHAKD 550
Db 559 KOKRTALHLASANGNSEVVKLLLDRCQLNVLDNKKRTALIKAVOCQDECALMLEHGT 618
QY 551 DPTKKNRDGNTPLDLVKDGTDIQDLRGAALLDAKKGCLARVKKLSSPDNVNCRDTQ 610
Db 619 DPNIPDEYN----- 628
QY 611 GRHSTPLHLAAGNRYKIVQLLLO-HGADVHAKDGLVPLHNACSYGHYDVAAALLIKYNA 670
Db 629 -----TTLHYAIYNEDKLMAKALLYGADIESKNKHLGTPLLGVHGEQKQVVKFLIKKKA 684
QY 671 CVNATDKWAFPLHFAAGNRYKIVQLLLO-HGADVHAKDGLVPLHNACSYGHYDVAAALLIKYNA 730
Db 685 NLNALDRYGRITALILAVCCGSAIVSLLLEQNDIVSSODLSQOTAREYAVSSHVHVI--- 741
QY 731 AMPPSALPSCIYKPOVLINGVRSPGATADALSSGSPSPSSLSAASSLDNLSGFSSELSVVS 790
Db 742 -----COLLSDYKEKQMLKTSF-----NSNPEQDLKLTSEESQRLGSENGQPELENFM 781
QY 791 SSGTEGASSLEKKEVPGVDFTSQFVRLNGLHLEMDIFEREQITLDVLVEMGHKELKEIG 850
Db 782 -----ENSQPEKMSQPEINKD-----GDRE-----VE-----BEMKKHE 811
QY 851 INAYGHRHLKIGVERLISGOQGLNPLYTLNTSGSTILIDLSPDDKEFQVEEEMOSTV 910
Db 812 SNNVGLLENITNGV-TAGNGDGLIPQKRSRT-----PENQOFPDNESEYHRI 859
QY 911 REHRDGGHAGGIFNRYNLIKTKQVCK-----KLWERYTHRRKEYSEENHNH----- 957
Db 860 CE-----LVSDYKEKMPKYSSENSPEQDLKLTSEESQRLGSENGQPELENFM 910
QY 958 ANERMLFHGSPFVNAIHKGFDE 980
Db 911 AIEEMKKHGS-----THVGPE 927

RESULT 13
US-09-822-827-378
; Sequence 378, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 378
; LENGTH: 1719
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-822-827-378
Query Match 6.6%; Score 379.5; DB 10; Length 1719;
Best Local Similarity 20.9%; Pred. No. 5.6e-19;

Search completed: February 12, 2003, 07:52:59
Job time : 27.5641 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 03:07:14 : Search time 21.6239 Seconds
(without alignments)
4890.316 Million cell updates/sec

Title: US-09-843-159b-3

Perfect score: 5769

Sequence: 1 GFRKDVVEYLLQNGASVQA.....AYPEYLITYQIMRDEGMVDG 1100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	877.5	15.2	1765	2 T42714	ankyrin 3, splice
2	877.5	15.2	1961	2 T42716	ankyrin 3, splice
3	875.5	15.2	4377	2 A55575	ankyrin 3, long sp
4	873.5	15.1	1943	2 T42713	ankyrin 3, splice
5	871.5	15.1	1940	2 T42715	ankyrin 3, splice
6	860.5	14.9	3924	2 S37431	ankyrin 2, neuroma
7	844	14.6	1856	2 B35049	ankyrin 1, erythro
8	844	14.6	1880	2 A35049	ankyrin 1, erythro
9	842	14.6	1881	1 SJHUK	ankyrin 1, erythro
10	839.5	14.6	1848	2 S37771	ankyrin, erythrocy
11	837.5	14.5	1862	2 T49502	ankyrin - mouse
12	834	14.5	1549	2 T13940	ankyrin - fruit fl
13	814.5	14.1	2039	2 T15347	ankyrin-related un
14	628	10.9	791	2 T42691	hypothetical prote
15	545.5	9.5	1411	2 S30355	alpha-latroinsecto
16	514	8.9	934	1 H71274	probable ankyrin -
17	496.5	8.6	1062	2 T30255	inversin - mouse
18	491.5	8.5	1062	2 T14151	Inv protein - mus
19	464.5	8.1	1401	2 S11527	alpha-latrotoxin p
20	456.5	7.9	1188	2 T19552	hypothetical prote
21	451	7.8	1058	2 D82654	ankyrin-like prote
22	431.5	7.5	426	2 AE2149	hypothetical prote
23	426	7.4	991	2 T25412	hypothetical prote
24	405.5	7.0	1184	2 T00253	gene Ankhzn protel
25	388	6.7	397	2 T46445	hypothetical prote
26	372.5	6.5	2584	2 T24158	hypothetical prote
27	372.5	6.5	2806	2 T24157	hypothetical prote
28	363.5	6.3	1031	2 T43458	hypothetical prote
29	350	6.1	557	2 T46507	hypothetical prote

ALIGNMENTS

RESULT 1

T42714

ankyrin 3, splice form 2 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42714

R:Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo

J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin ge the repeat domain.

A:Reference number: Z22237; MUID:95340633; PMID:7615634

A:Accession: T42714

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-1765 <PE>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605.1

A:Experimental source: strain C57Bl/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 1587/1

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

Query Match 15.2%; Score 877.5; DB 2; Length 1765;
Best Local Similarity 30.9%; Pred. No. 1.8e-46;
Matches 277; Conservative 125; Mismatches 377; Indels 117; Gaps 22;

QY 3 GRKDVVEYLLQNGASVQARDGGLPLHNACSFGEAEVYNLLRHGADPNARDNNYTPPL 62

Db 68 GHVEVVSSELLQREANVDAATKKGNTALHISLAGQAEVVKVLVTNGANVNAQSQNGFTPL 127

QY 63 HEAAIKGIDVCIVLLQHGAEPTIRNTDGRALDIA-----DPSAKAVLTGEYK---KDE 114

Db 128 YMAAGENHLEVVRFLLDNGASQSLATEDGFTPLAYALQQHQDQVVSLLLENDITKGKVLRLP 187

QY 115 LLESARSNEEKMMALLPLPNVNCASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK 174

Db 188 ALHIAARKDDTKAAALLQNDTNADVESKSGFTPLHIAAHYGNINVTLLNRAAANDFT 247

QY 175 DKGLDVLPLHNACSYGHEVTELLVVKHGACVNAMDILQWFTPLHEAASKNRVESCILLSYG 234

Db 248 ARNDITPLHVASKRGNNANVKKLLDRGAKIDAKTRDGLTPLHCGARSHEGVVEMLLDRS 307

QY 235 ADPTLLNCHNKSAIDIAPTPQLKRLAYEFKSHLSLQARADVTPIKKHLSLEWNVFKH 294

Db 308 A-PILSKTKN-----GLSPLHMTQGD-----HLNCVQLLLOH 339

QY 295 -----PQTHETALHCAASPYPKRKOICELLRLKGANINEKTEFTPLHVASEKAHN 347

Db 340 NVFVDVDTNDYLTALHVAACHGCHKY---VAKVLLDKKSPNAKALNGFTPLHIACKKNRI 396

hypothetical prote
death-associated p
hypothetical prote
hypothetical prote
hypothetical prote
myosin heavy chain
2-5A-dependent RNA
ankyrin related pr
hypothetical prote
probable ankyrin
transcription fact
myosin-light-chain
sex-determining pr
hypothetical prote
probable membrane
related to 26s pro

QY 348 DVVEVVVKHEAKYNALDNLGQSLTHRAAYGCHLQTCRLLSYGCDDPNIIISLOGFTALOMG 407
 Db 397 RVMEILLKHGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHGASPNVTNVRGETALHMA 456
 QY 408 -----NENVOQLLOEGISLGNSEADQ--LLEAAKAGDVETVKKLCTV-OSVNCRDIEGR 459
 Db 457 ARSGQAEVVRVLYVDGGAQVEAKAKDDQTPHISARLGKADIVQOQLQOOGASPNAAATTS- 515
 QY 460 QSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGLVPLHNACSYGHEVAELLVKHCAYV 519
 Db 516 -YTPLHAAAREGHEDVAFFLLDHGASLSITTKGFTPLHVAAYGKLEVASLLLOKSASP 574
 QY 520 NVADLWKFTPLHPEAAKGYEICKLLOHGADPTKKNRDNCTPLDL-VKDGDDTDIQLL- 577
 Db 575 DAAGKSGLTPLHVAAYHYNQKVALLLDQOGASPHAAAKNGYTPHIAAKKNQMDIATSL 634
 QY 578 -----RGDAALLDAAKKGCLARVKKLSSPD-NVNCRTQGRHSTPLHLAAGYNN 625
 Db 635 EYGADANAVTROGIAVSHVLAAGQGHVDMVSLLSRANVNLNKS- -LTPLHAAQEDR 692
 QY 626 LEVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFPLHE 685
 Db 693 VNVAEVLVNOGAHVDAQTKMGYTPLVHGYGNKIKVNFLLQHSKVNAKTKNGYALHQ 752
 QY 686 AAKGRTQCALLAHGAADPTLKNQEGQTPDLVS-----ADDYSALITAAAPPSPALP 738
 Db 753 AAQOGHTHIINVLQNNASPNELTVNGNTALATARRLGIVSVVDTLKVVTEIIMTTTTIT 812
 QY 739 SCVK-----POVLNCGVRSPGATADALS- -GPFSELSSVWSSSGT 794
 Db 813 EXHKMNPETMNEVLD-----MSDDEVKASAPKLSGDEY-----ISDGE 853
 QY 795 EGASSLEKKEVPGVDFSIQFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIG 850
 Db 854 EGDKCTWEK-IPKQVEVLK-----SEDAITGDTDKYLGPDQLKELG 894
 RESULT 2
 T42716
 ankyrin 3, splice form 4 - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
 C:Accession: T42716
 R:Peters, L.B.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, I.
 J. Cell Biol. 130, 313-330, 1995
 A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
 the repeat domain.
 A:Reference number: 222237; MUID:95340633; PMID:7615634
 A:Accession: T42716
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1961 <PET>
 A:Cross-references: EMBL:L40632; NID:g710548; PID:g710552; PIDN:AAB01607.1
 A:Experimental source: strain C57BL/6J; kidney
 C:Genetics:
 A:Gene: Ank3
 A:Map position: 10
 C:Superfamily: ankyrin; ankyrin repeat homology
 C:Keywords: alternative splicing

Query Match 15.2%; Score 877.5; DB 2; Length 1961;
 Best Local Similarity 30.9%; Pred. No. 2.1e-46;
 Matches 277; Conservative 125; Mismatches 377; Indels 117; Gaps 22;

QY 3 GRKDVVEYLLQNGASVQARDGGLIPLHNACSFHAEVYNLLLRHGADPNARDNNWNYTPL 62
 Db 68 GIVEVVELLQREANVDRAATKKGNTALHTASLAGQAEVVKVLVTNGANNVNAOSQNGFTPL 127
 QY 63 HPAALIKGIDVCIVLLOHGAETPIRNTQRTALDLA-----DPSAKAVITGEYK---KDE 114
 Db 128 YMAAENHLEVVRELLDNGASQSLATEDGFTPLAVALQOCHDQVVSLLLENTKGVRLP 187
 QY 115 LLESARSNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLLOHGADVHAK 174

Db 188 ALHIAARKDDTKAAALLQNDTNADVESKSGFTPLHIAAHYGNINVAITLLNRAAAVDFT 247
 QY 175 DKGDVPLHNACSYGHEVTELLVKGACVYAMDLWQFTPLHHEAASKNRVEYCSLLLSYG 234
 Db 248 ARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARGSHGEOVEMLLDRS 307
 QY 235 ADPTLLNCHNKSAIDLAAPTQPKERLAYEFKSHLSLOAAREADVTRIKKHLSLEMVNFKH 294
 Db 308 A-PILSKTKN-----GLSPLHMAATQGD-----HLNCVQLLQHQ 339
 QY 295 -----PQTHETALHCAASPYKPKQICELLRLRGANINEKTEFLTPLHVAASEKAHN 347
 Db 340 NVPVDDVTNDYLTALHVAACHGHYK---VAKVLLDKKASPNALANGFTPLHIAACKNRI 396
 QY 348 DVVEVVVKHEAKYNALDNLGQSLTHRAAYGCHLQTCRLLSYGCDDPNIIISLOGFTALOMG 407
 Db 397 RVMEILLKHGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHGASPNVTNVRGETALHMA 456
 QY 408 -----NENVOQLLOEGISLGNSEADQ--LLEAAKAGDVETVKKLCTV-OSVNCRDIEGR 459
 Db 457 ARSGQAEVVRVLYVDGGAQVEAKAKDDQTPHISARLGKADIVQOQLQOOGASPNAAATTS- 515
 QY 460 QSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGLVPLHNACSYGHEVAELLVKHCAYV 519
 Db 516 -YTPLHAAAREGHEDVAFFLLDHGASLSITTKGFTPLHVAAYGKLEVASLLLOKSASP 574
 QY 520 NVADLWKFTPLHPEAAKGYEICKLLOHGADPTKKNRDNCTPLDL-VKDGDDTDIQLL- 577
 Db 575 DAAGKSGLTPLHVAAYHYNQKVALLLDQOGASPHAAAKNGYTPHIAAKKNQMDIATSL 634
 QY 578 -----RGDAALLDAAKKGCLARVKKLSSPD-NVNCRTQGRHSTPLHLAAGYNN 625
 Db 635 EYGADANAVTROGIAVSHVLAAGQGHVDMVSLLSRANVNLNKS- -LTPLHAAQEDR 692
 QY 626 LEVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFPLHE 685
 Db 693 VNVAEVLVNOGAHVDAQTKMGYTPLVHGYGNKIKVNFLLQHSKVNAKTKNGYALHQ 752
 QY 686 AAKGRTQCALLAHGAADPTLKNQEGQTPDLVS-----ADDYSALITAAAPPSPALP 738
 Db 753 AAQOGHTHIINVLQNNASPNELTVNGNTALATARRLGIVSVVDTLKVVTEIIMTTTTIT 812
 QY 739 SCVK-----POVLNCGVRSPGATADALS- -GPFSELSSVWSSSGT 794
 Db 813 EXHKMNPETMNEVLD-----MSDDEVKASAPKLSGDEY-----ISDGE 853
 QY 795 EGASSLEKKEVPGVDFSIQFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIG 850
 Db 854 EGDKCTWEK-IPKQVEVLK-----SEDAITGDTDKYLGPDQLKELG 894
 RESULT 3
 A55575
 ankyrin 3, long splice form - human
 N:Alternate names: ankyrin G
 C:Species: Homo sapiens (man)
 C>Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 20-Sep-1999
 C:Accession: A55575
 R:Kordeli, E.; Lambert, S.; Bennett, V.
 J. Biol. Chem. 270, 2352-2359, 1995
 A:Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the
 A:Reference number: A55575; MUID:95138209; PMID:7836469
 A:Accession: A55575
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-4377 <KOR>
 A:Cross-references: GB:U13616; NID:g608024; PIDN:AAA64834.1; PID:g608025
 C:Genetics:
 A:Gene: GDB:ANK3
 A:Cross-references: GDB:424503; OMIM:600465
 A:Map position: 10q21-10q21
 C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homo

C:Keywords: alternative splicing; peripheral membrane protein

F:73-105/Domain: ankyrin repeat homology <AN01>
 F:106-138/Domain: ankyrin repeat homology <AN02>
 F:139-171/Domain: ankyrin repeat homology <AN03>
 F:172-200/Domain: ankyrin repeat homology <AN04>
 F:201-233/Domain: ankyrin repeat homology <AN05>
 F:234-266/Domain: ankyrin repeat homology <AN06>
 F:267-299/Domain: ankyrin repeat homology <AN07>
 F:300-332/Domain: ankyrin repeat homology <AN08>
 F:333-365/Domain: ankyrin repeat homology <AN09>
 F:366-398/Domain: ankyrin repeat homology <AN10>
 F:399-431/Domain: ankyrin repeat homology <AN11>
 F:432-464/Domain: ankyrin repeat homology <AN12>
 F:465-497/Domain: ankyrin repeat homology <AN13>
 F:498-530/Domain: ankyrin repeat homology <AN14>
 F:531-563/Domain: ankyrin repeat homology <AN15>
 F:564-596/Domain: ankyrin repeat homology <AN16>
 F:597-629/Domain: ankyrin repeat homology <AN17>
 F:630-662/Domain: ankyrin repeat homology <AN18>
 F:663-695/Domain: ankyrin repeat homology <AN19>
 F:696-728/Domain: ankyrin repeat homology <AN20>
 F:729-761/Domain: ankyrin repeat homology <AN21>
 F:762-794/Domain: ankyrin repeat homology <AN22>
 F:795-827/Domain: ankyrin repeat homology <AN23>

Query Match 15.2%; Score 875.5; DB 2; Length 4377;
 Best Local Similarity 30.5%; Pred. No. 1e-45;
 Matches 273; Conservative 126; Mismatches 362; Indels 135; Gaps 21;

Qy 3 GRKDVVEYLLQNGASQVARDGGLIPLHNACSFHAEVNVNLLLRHGADPNARDNNWYTP 62
 Db 85 GHVEVVSSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLTNGANVNAQSQNGFTPL 144

Qy 63 HEAAIKGKIDVICVILQHGAEPTIRNTDGTALDIA-----DPSAKAVLTGEYK---KDE 114
 Db 145 YMAQENHLEVVKLLDNGASQSLATEDGFTPLAVALQGHQDVVSLLENDTKGVRLP 204

Qy 115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHAAAGYNRKIVTQVLLQHGADVHAK 174
 Db 205 ALHIAARKDDTKAAALLQNDNNAVDSEKSGFTPLHIAAHYGNINVTALLNRAAAVDFT 264

Qy 175 DKGDVPLHNACSYGHEVTELLVKHGACVNAWMDLWQFTPLHAEAKNREVCVCSLLSYG 234
 Db 265 ARNDITPLHVAKRGNANVMKLLLDGAKIDAKTRDGLTPLHCGARSGHEQVVEMLLRA 324

Qy 235 ADPTLLNCHNKSADLAPTLPOLKERLAYEFKGHSLQAAAREADVTRIKKHLSLEWVNFKH 294
 Db 325 A-PILSKTKN-----GLSPHLMATQGD-----HLNCVQLLQHQ 356

Qy 295 -----POTHEALTHCAASPYPKRKQICELLLRKGANINEKTEFTPLPHVASEKAHN 347
 Db 357 NVPVDDVTNDYLTALHVAACHGHYK---VAKVLLDKKPNKAKALNGFTPLHIAACKNRI 413

Qy 348 DVVEVVVKHAKVNALDNLGQTSLEHRAAYCGHLOTCRLLSYGCDPNIISLOGFTALQMG 407
 Db 414 KVMELLKHGASIQAVTESGLTPIHVAAPFMGHVNIYSQMLHGAASPNTINVRGETALHWA 473

Qy 408 -----NENVOQLQEGISGNSEADRO--LLEAAKAGDVETVKKLCVTV-QSVNCRDIEGR 459
 Db 474 ARSQAEVRYLVQDQAGVQAEAKADDOFTPLHISARLGKADIVQQLQOOGASPNAAITSG- 532

Qy 460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKHGAUV 519
 Db 533 -YTPLHLSAREGHEDVAAFLDLDGASLSITTKGFTPLHVAARYKGLVEANLLLOKASAP 591

Qy 520 NVADLWKFETPLHEAAKKGKYEICKLLLOHGADPTKKNRDGNTPDL-LVKDGDFTDIQDL- 577
 Db 592 DAAGKSLTPLHVAHYDQKVAILLDDQGASPHAAAKNGYTPPLHIAAKNQMDIATLTL 651

Qy 578 -----RGDAALLDAAKGCLARVKKLSSPD--NVNCRDTQGRHSTPLHAAAGYNN 625
 Db 652 EYGADANAVTROGSIASVHAAQEGHVDVNSLLLRNANVNLNSKSG--LTPLHAAQEDR 709

Qy 626 LEVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYVNAACVNAATDKWAFPLHE 585
 Db 710 VNVAEVLVNOGAHVDAQTKMGYTPPLHVGCHYGIKIVNFLLQHSKAVNAKTNGYTPLHQ 769

Qy 686 AAKGRTQICALLAHGADPTLKNQEGQTPDL-----VSADDSALTLT-AAAPPSALP 738
 Db 770 AAOQGHTHIIVLQNNASPNELTVNGNTALGIARRLGYSIVVDTLKIVTEETMTTITVT 829

Qy 739 SCYK---POVLNGVRSPGATADALSSGSPSSLSAASSLDNLGSGFSSELSSVSSSGTE 795
 Db 830 EKHKMVPETMNEVLD-----MSDDEVVRKANAPEMLSDG----- 863

Qy 796 GASLEKKEVPGVDFSTIQVRNGLGLEHLMDFEREQ-ITLDVLVEMGHKELKEIG 850
 Db 864 -----EYISDVBEGEDAMTGDTKYLGPDQLKELG 893

RESULT 4
 T42713
 ankyrin 3, splice form 1 - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #Text_change 04-Mar-2000
 A:Accession: T42713
 R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, M.; Turtzo
 J. Cell Biol. 130, 313-330, 1995
 A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene repeat domain.
 A:Reference number: Z22237; MUID:95340633; PMID:7615634
 A:Accession: T42713
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1943 <PEP>
 A:Cross-references: EMBL:L40632; NID:g710548; PID:g710550; PIDN:AA801606.1
 A:Experimental source: strain C57BL/6J; Kidney
 C:Genetics:
 A:Gene: Ank3
 A:Map position: 10
 A:Introns: 855/1
 C:Function:
 A:Description: supposed to play an important role in the polarized distribution of ma
 A:Note: major kidney ankyrin
 C:Superfamily: ankyrin; ankyrin repeat homology
 C:Keywords: alternative splicing

Query Match 15.1%; Score 873.5; DB 2; Length 1943;
 Best Local Similarity 30.6%; Pred. No. 3.7e-46;
 Matches 274; Conservative 122; Mismatches 365; Indels 135; Gaps 21;

Qy 3 GRKDVVEYLLQNGASQVARDGGLIPLHNACSFHAEVNVNLLLRHGADPNARDNNWYTP 62
 Db 68 GHVEVVSSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLTNGANVNAQSQNGFTPL 127

Qy 63 HEAAIKGKIDVICVILQHGAEPTIRNTDGTALDIA-----DPSAKAVLTGEYK---KDE 114
 Db 128 YMAQENHLEVVKLLDNGASQSLATEDGFTPLAVALQGHQDVVSLLENDTKGVRLP 187

Qy 115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHAAAGYNRKIVTQVLLQHGADVHAK 174
 Db 188 ALHIAARKDDTKAAALLQNDNNAVDSEKSGFTPLHIAAHYGNINVTALLNRAAAVDFT 247

Qy 175 DKGDVPLHNACSYGHEVTELLVKHGACVNAWMDLWQFTPLHAEAKNREVCVCSLLSYG 234
 Db 248 ARNDITPLHVAKRGNANVMKLLLDGAKIDAKTRDGLTPLHCGARSGHEQVVEMLLDRS 307

Qy 235 ADPTLLNCHNKSADLAPTLPOLKERLAYEFKGHSLQAAAREADVTRIKKHLSLEWVNFKH 294
 Db 308 A-PILSKTKN-----GLSPHLMATQGD-----HLNCVQLLQHQ 339

Qy 295 -----POTHEALTHCAASPYPKRKQICELLLRKGANINEKTEFTPLPHVASEKAHN 347
 Db 340 NVPVDDVTNDYLTALHVAACHGHYK---VAKVLLDKKPNKAKALNGFTPLHIAACKNRI 396

Qy 348 DVVEVVVKHAKVNALDNLGQTSLEHRAAYCGHLOTCRLLSYGCDPNIISLOGFTALQMG 407


```

Db 397 RYMEILLKHGASIQAVTESGLTPIHVAFAFMGHVNIQSOLMHGASPNNTNVRGETALHMA 456
QY 408 -----NENVOQLLOEGISLGNSEADRO--LLEAAKAGDVETVKKLCIV-OSVNCRDIEGR 459
Db 457 ARSGOAEVVRVYLDQGAQVEAKAKDDQTPHISARLGKADIVQOOLQOOGASPNAAATTS- 515
QY 460 QSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGGLVPLHNAACSYGHYEVAELLYKHGAVV 519
Db 516 -YTPLHLAAREGHEDVAFAFLDHGASLSITTKGFTPLHVAAKYKGLVEASLLLOKQAS 574
QY 520 NVADLWKFPTPLHEAAKGYEICKLLOHGADPTKKNRDGNTPLDL-VKGDGTDIQLL- 577
Db 575 DAAGKSLGTLPLHVAHYDNQVALLLLDQGASPHAAAKNGYTPHIAAKKNQMDIATSL 634
QY 578 -----RGDAALLDAAKGCCLARVKKLSPD-NVNCRDTOGRHSTPLHAAAGYNN 625
Db 635 EYGADANAVTQGTASVHLAAQEGHDMVLSLLSRNANVLSNKG--LTPHLHAAQEDR 692
QY 626 LEVAEYLLOHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWFTPLHE 685
Db 693 VNVAEVLVNOGAHVDAQTKMGYTPLVHVGCHYGNIKIVNFFLOHSAKYNKNGYTAHQ 752
QY 686 AAKGRTOLCALLAHGADPTLKNQEGTPLDIVS-----ADDVSALLTAAMPSPALP 738
Db 753 AAQOGHTHIINVLQNNASPNELTVNGTALAIARRLGYSVVDTLKVVTTEEIMTTTIT 812
QY 739 SCYK---POVLNGVRSPGATADALSSPSSPSLSAASSLDNLS-GSFSELSVVSSSGT 794
Db 813 EKHKNVPETMNEVLN-----MSDDEVKASAEKLSDEY-----ISDGE 853
QY 795 EGASSLEKKEVPGVDFISITQFVRLNGLHLMDFEREOITLDVLVEMGHKELKEIG 850
Db 854 EG-----EDAITGDTDKYLGPODLKELG 876

RESULT 5
T42715
ankyrin 3, splice form 3 - mouse
C:Species: Mus musculus (house mouse)
C:date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42715
R:Peters, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Vialamas, M.; Turlizo, I.
J. Cell Biol. 130, 313-330, 1995
A:title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A:Reference number: 222237; MUID:95340633; PMID:7615634
A:Accession: T42715
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1940 <PEP>
A:Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDN:AA01604.1
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
A:Introns: 834/1
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

```

Query Match 15.1%; Score 871.5; Db 2; Length 1940;
 Best Local Similarity 28.3%; Pred. No. 5e-46;
 Matches 320; Conservative 159; Mismatches 464; Indels 189; Gaps 35;

```

QY 3 GRKDVVEYLLQNGASVOARDGGLIPLHNAACSYGHAEVYNLLIRHGADPNARDNNMYTPL 62
Db 68 GHVEYVEYLLQNGASVOARDGGLIPLHNAACSYGHAEVYNLLIRHGADPNARDNNMYTPL 62
QY 63 HEAAATKGTIDVCIVLLOHCAEPTIRNTDRTALDLA-----DPSAKAVLTGEYK----KDE 114
Db 128 YMAAENHLEVVRFLLDNGASLSATGDTGFTPLAVALQOQCHDQVVSLLLENDTKGKVRLP 187
QY 115 LLESARSGNEERKWMALLTPLNVNCHASDGRKSTPLHAAAGYNRVKIVQLLLOHGADVHAK 174

```

```

Db 188 ALHIAARDDTKAAALLQNDTNADVESKSGFTPLHIAAHYGNINVTALLNRAAAVDFT 247
QY 175 DKGDVPLHNAACSYGHYEVTETLLVKKHGCNVAMDLWQFPTPLHEAAKKNRVEVCSLLSYG 234
Db 248 ARNDITPLHVASKRGNANWVKILLDRGAKIDAKTRDGLTPLHCGARSHEQVVEMLLDRS 307
QY 235 ADPTLLNCHNKAIDLAPTQPKERLAYEFYKSHLSLQAAAREADVTRIKKHLSLEMVNFKH 294
Db 308 A-PILSKTKN-----GLSPLHMAQGD-----HLNCVQLLLOH 339
QY 295 -----PQTHETALHCAASPYPKRKOICELLRLKGNINEKTEFTPLPLHVAASEKAHN 347
Db 340 NVPVDDVTNDYLTALHVAACHYK---VAKVLDKKASNAKALNGFTPLHIACKKRI 396
QY 348 DVVEVVKHEAKVNALDNLGQTSLHRAAYCGHLOTCRLLSYGCDPNIISLQGTALOMG 407
Db 397 RVMEILLKHGASIQAVTESGLTPIHVAFAFMGHVNIQSOLMHGASPNNTNVRGETALHMA 456
QY 408 -----NENVOQLLOEGISLGNSEADRO--LLEAAKAGDVETVKKLCIV-OSVNCRDIEGR 459
Db 457 ARSGOAEVVRVYLDQGAQVEAKAKDDQTPHISARLGKADIVQOOLQOOGASPNAAATTS- 515
QY 460 QSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGGLVPLHNAACSYGHYEVAELLYKHGAVV 519
Db 516 -YTPLHLAAREGHEDVAFAFLDHGASLSITTKGFTPLHVAAKYKGLVEASLLLOKQAS 574
QY 520 NVADLWKFPTPLHEAAKGYEICKLLOHGADPTKKNRDGNTPLDL-VKGDGTDIQLL- 577
Db 575 DAAGKSLGTLPLHVAHYDNQVALLLLDQGASPHAAAKNGYTPHIAAKKNQMDIATSL 634
QY 578 -----RGDAALLDAAKGCCLARVKKLSPD-NVNCRDTOGRHSTPLHAAAGYNN 625
Db 635 EYGADANAVTQGTASVHLAAQEGHDMVLSLLSRNANVLSNKG--LTPHLHAAQEDR 692
QY 626 LEVAEYLLOHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWFTPLHE 685
Db 693 VNVAEVLVNOGAHVDAQTKMGYTPLVHVGCHYGNIKIVNFFLOHSAKYNKNGYTAHQ 752
QY 686 AAKGRTOLCALLAHGADPTLKNQEGTPLDIVS-----ADDVSALLTAAMPSPALP 738
Db 753 AAQOGHTHIINVLQNNASPNELTVNGTALAIARRLGYSVVDTLKVVTTEEIMTTTIT 812
QY 739 SCYK---POVLNGVRSPG-----ATADALSS-----GP----- 763
Db 813 EKHKNVPETMNEVLN-----MSDDEVKASAEKLSDEY-----ISDGE 853
QY 795 EGASSLEKKEVPGVDFISITQFVRLNGLHLMDFEREOITLDVLVEMGHKELKEIG 850
Db 854 EG-----EDAITGDTDKYLGPODLKELG 876

RESULT 6
S37431
ankyrin 2, neuronal long splice form - human
N:Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythro
N:Contains: ankyrin 2, short form

```

C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
C:Accession: S37431; A39643; B39643; A40334; A49462; S14533; S14569
R:Chan, W.
submitted to the EMBL Data Library, September 1993
A:Reference number: S37431
A:Accession: S37431
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-3924 <CHA>
A:Cross-references: EMBL:Z26634; NID:9406287; PIDN:CAA81387.1; PID:9406288
R:Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.
J. Cell Biol. 114, 241-253, 1991
A:Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a
A:Reference number: A39643; MUID:91302466; PMID:18330053
A:Accession: A39643
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2077 <Orl>
A:Cross-references: GB:X56957
A:Accession: B39643
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1443,3585-3924 <OTT>
A:Cross-references: EMBL:X56958
R:Tse, W.T.; Menniger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Ward,
Genomics 10, 858-866, 1991
A:Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.
A:Reference number: A40334; MUID:92009921; PMID:1833308
A:Accession: A40334
A:Molecule type: DNA
A:Residues: 463-474, 'PE', 477-495 <TSE>
A:Cross-references: GB:M37123; NID:9178647; PIDN:AAA62828.1; PID:9178648
R:Chan, W.; Kordeili, E.; Bennett, V.
J. Cell Biol. 123, 1463-1473, 1993
A:Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and se
A:Reference number: A49462; MUID:94075409; PMID:8253844
A:Accession: A49462
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-3924 <RES>
A:Cross-references: EMBL:Z26634; NID:9406287; PIDN:CAA81387.1; PID:9406288
C:Genetics:
A:Gene: GDB:ANK2
A:Cross-references: GDB:127607; OMIM:106410
A:Map position: 4q25-4q27
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:2-3924/Product: ankyrin 2, long form #status predicted <MAT>
F:2-1443,3585-3924/Product: ankyrin 2, short form #status predicted <MA2>
F:63-95/Domain: ankyrin repeat homology <AN01>
F:96-128/Domain: ankyrin repeat homology <AN02>
F:129-161/Domain: ankyrin repeat homology <AN03>
F:162-190/Domain: ankyrin repeat homology <AN04>
F:191-223/Domain: ankyrin repeat homology <AN05>
F:232-264/Domain: ankyrin repeat homology <AN06>
F:265-297/Domain: ankyrin repeat homology <AN07>
F:298-330/Domain: ankyrin repeat homology <AN08>
F:331-363/Domain: ankyrin repeat homology <AN09>
F:364-396/Domain: ankyrin repeat homology <AN10>
F:397-429/Domain: ankyrin repeat homology <AN11>
F:430-462/Domain: ankyrin repeat homology <AN12>
F:463-495/Domain: ankyrin repeat homology <AN13>
F:496-528/Domain: ankyrin repeat homology <AN14>
F:529-561/Domain: ankyrin repeat homology <AN15>
F:562-594/Domain: ankyrin repeat homology <AN16>
F:595-627/Domain: ankyrin repeat homology <AN17>
F:628-660/Domain: ankyrin repeat homology <AN18>
F:661-693/Domain: ankyrin repeat homology <AN19>
F:694-726/Domain: ankyrin repeat homology <AN20>
F:727-759/Domain: ankyrin repeat homology <AN21>
F:760-792/Domain: ankyrin repeat homology <AN22>
F:793-825/Domain: ankyrin repeat homology <AN23>

Query Match 14.9%; Score 860.5; DB 2; Length 3924;
Best Local Similarity 29.1%; Pred. No. 7.5e-45;
Matches 297; Conservative 147; Mismatches 404; Indels 174; Gaps 29;
Qy 3 GRKDVEYLLONGASVOARDGGGLIPLHNACSGHAEVWNLRLRHGADPNARDNWNWYPL 62
Db 75 GHVGLVQELLGRSSVDSATKNTALHIALASLAGQAEVWVVLKREGANINAQSONGFTPL 134
Qy 63 HEAAIKGKIDVICVILLOHGAFTIRNTDGT---ALDLADPSAKAVLTGEYKDEL--- 115
Db 135 YMAOENHIDVVYKLLLENGANOSTATEDGFTPLAVALQQHGNQAVALLLENDTKGKVRP 194
Qy 116 -LESARSGNEEKMMALLTPLNVNCHASDGRKS-----TPHLAAGYNRVKIVQL 163
Db 195 ALHIAARKDDTKSAALLLQ---NUHNADVQSKMNVNRTTESGFTPLHIAAHYGNVATL 251
Qy 164 LLQHGADVHAOKGDLVPLHNACSYGHEVTELLVKGACVNMALWQFTPLHEAASKNR 223
Db 252 LLNRGAADVFTARNGITPLHVASKRGNTNMVKLLLDGRGQIDAKTRDGLTPLHCAARSGH 311
Qy 224 VEVCSLLLSYGADPTLLNCHNKSALDLAPTQPKERLAYEFKSHSLLOAAAREADVTRIKK 283
Db 312 DOVVELLLERGA-PLLAARTKN-----GLSPLHMAAQGDHVECVK 349
Qy 284 HLSLEMVNFKHP-----QTHETALHCAASPYPRKQICELLLRKGANINEKTEFTPL 338
Db 350 HL----LQHKAPVDVDTLDYLTALHVAA---HCGHYRVTKLLDKRANPNARALNGFTPL 402
Qy 339 HVASEKAHNDVVVVVVKHEAKVNALDNLGQTSLHRAAYCGHLOTCRLLSLLSGCDPNISL 398
Db 403 HIACKNKRIKVMELLKYGASIQAITESGLTPIHVAAPFMGHLNIVLLLLONGASPDVTNI 462
Qy 399 OGFTALQMGNEVQQLQEGISLGNSEADROLLEAAKAGDVETVKYKLCVQSVNCRDIEG 458
Db 463 RETALHM-----AARAGQVEVVR---CLLRNGALVDARA 494
Qy 459 R-QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDGGGLVPLHNACSYGHEVTELLVKGHA 517
Db 495 REEQTPHIAASRLGKTEIVQLLQHMHPDAATTNGYTPHISAREGOVDVASVLEAGA 554
Qy 518 VVNVADLWKFPLHFAAAKGYEICKLLLOHCADPTTKNRDGNTPDLVKGDDTDIDOLL 577
Db 555 AHSLATKKGFTPLHVAAYKGLDVAKLLQRRRAADSAGKNGLTPLHVAHYDN----- 608
Qy 578 RGDAAALLDAAKGLARVKKLSPPDNVNCRTDQGRHSTPLHLAAGYNNLEVAEYLQHGA 637
Db 609 -----QKVALLLLEKASP-----HATAKNGYTPHIAAKKNQMIATLLNYGA 653
Qy 638 DYNAQDKGLIPLHNAASGHVDVVAALLIKYNACVNATDKNAFTPLHEAAQKGRTOCAL 697
Db 654 ETNIVTKQGVTPHLIASQEGHDMVTLLLDKGANIHMSTKSGTSLHLAAQEDKYNVADI 713
Qy 698 LLAHGADPTLKNOEGOTPLDLIV---SADDVSALLTAAMPSPA-LPSCVKP----- 743
Db 714 LTKHGADQDAHTKGLTTPHIVACHYGNVKNVWFLKQGANVNAKTKNGYTPHQAQOQH 773
Qy 744 -QVLNGVRSPGATADALSSGSPSSLSAASLDNLGSSFSLSVSSSGTEGASLLEK 802
Db 774 THINVLLOHGAKPNTVA--NGNTALAIKRLGYI--SVVDTLKVVTVEVTTTTTITE 829
Qy 803 KEVPGVDFSIQTQVRNMLGHEHMDIFERE---QITLDVLVEMGHKELKEIGNAYCHRRH 859
Db 830 KHKLNVPETMTE-----VLDVSDERGGDDTMTGDGGEYLRPEDLKLSDGDS----- 874
Qy 860 LIKGVRLISGO--QGLNPLYLTNTSGSTILIDILSPDKFQSVVEEQSVREHROGG 917
Db 875 -----LPSSQFLDGMN-ILYRLSLEGG-----RSDLSRFSSSDRSHTLSHASYLRD--- 918
Qy 918 HAGGIFNRYNLIKQKVCN-KKLWERYTHRRKEVSEENHNHANERMLFPGSPFVNAIHK 976
Db 919 --SAVMDSDSVIPSHQVSTLAKAEARNRYLSWGNTENLQNA-----LSSSP-----IHS 966

```

Qy 977 GF 978
Db 967 GF 968

RESULT 7
B35049
Ankyrin 1, erythrocyte splice form 3 - human
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
C:Species: Homo sapiens (man)
C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
C:Accession: B35049
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; K
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370; PMID:1689849
A:Accession: B35049
A:Status: preliminary
A:Molecule type: mRNA
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:2-1513,1676-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MA2>
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>
F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match 14.6%; Score 844; DB 2; Length 1856;
Best Local Similarity 30.1%; Pred. No. 2,5e-44;
Matches 242; Conservative 121; Mismatches 263; Indels 177; Gaps 21;

Qy 3 GRKDVVEYLLQNGASVOARDGGLPLHNACSFGEAEVNNLLRHGADPNARDNNVYPL 62
Db 89 GDEVVRELVNNGANVNAQSQKGTPLVMAAQNHLVVKFLENGANQNVATEDGFTPL 148
Qy 63 HEAAIKGKIDVCIVLQHGAEPTI-----RNTDGRITDLADPSAKAVLTGEYK 113
Db 149 AVAQOQHENVVVAHLNINYGKGVRLPALHIAARNDDTRTA-----AVLLQNDPNP 199
Qy 114 ELHESARSENEEKMMALLTPLVNVNCHASD-----GRKS-----TPHLAAGY 155
Db 200 DVL-SKTG-----FTPLHAAHYENLVNVAQLLNRGSSVNFPPQNGITPLHIASR 249
Qy 156 NRKIVVOLLQHGADVHAKDGLVPLHNACSYGHVEVTELLVKKHACVNAAMDLMQFTPL 215
Db 250 GNVIMVRLLDGRGAQITETKTDELTPHCAARNGHVRISEILLDHCAPIQAKTKNGLSPI 309

```

```

Qy 216 HEASKNREVEVCSLLLSYGADPTLLNCHNKSAIDLAPTLPOLKERLAYEFKHSLLQAARE 275
Db 310 HMAAGDHLDVCVRLLYQDAE-----ID----- 332
Qy 276 ADVTRIKKHLISLEMVNFKHPQTHETALHCAASPYPKRKOICELLRLKRGANINEKTKRFL 335
Db 333 -DIT-----LDHLTPHVA-----HGHIRVAKVLDKAGKPNRSALNGF 372
Qy 336 TPLHVASEKAHNDVVVVKHEAKVNALDNLGOTSLSHRAAYCGHQLQTCRLLLSYGCDPNI 395
Db 373 TPLHACKKNHVRVMEILLKTKGASIDAVTESGLTPLHVASEFMGHLPIVKNLIQRCASPNV 432
Qy 396 ISLOGFTALQMS-----NENVOQLLOE-----CISLGNSEADROLLE- 432
Db 433 SNKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPHCAARIGHTNMVKLLLEN 492
Qy 433 -----AAKAGDVETVKKLCQV--SVNCRDIEGRQSTPLHFAAGYNRS 474
Db 493 NANPNLATTAGTTPHIAARESHVETVLALLEKASQACMTKKG--FTPLHVAAYKGVKVR 550
Qy 475 VVEYLLQHGADVHAKDGLVPLHNACSYGHVEAEVLLVKKHGVNVNADLWK-FTPLHEA 533
Db 551 VAEULLERDAHPNAAKNGLTPLHVAVHNNLDIVKLLPRGSGPH-SPANNGVYTPHIA 609
Qy 534 AAKGYEICKLLQHGADPTKKNRDGNTPDL-VKDGDTDIDQLLRGDAALLDAKKKCL 592
Db 610 AKQOQVEVASLLQYGGSAANESVOGVTPLHIAAQEGHAEVALL----- 654
Qy 593 ARVKLLSPDNVNCRDTOGRHSTPLHIAAGYNNLEVAEYLLQHGADVNAODKGLIPLHN 652
Db 655 -----LSKQANGNLGKSG--LTPHLVVAQSGHVPVADVLKKGVMVDATTRMGYTPH 707
Qy 653 AASGYHVDVVAALLIKYNACVNAATDKWAFPLHEAAQKRGRTOLCALLAHGADPTLKNOEG 712
Db 708 ASHYGNIKLVKFLQHQADVNAATKLGYSPLHQAQOQHTDITVLLKNGASPNVSSDG 767
Qy 713 QTPLDL-----VSADVDVSALLT 729
Db 768 TTPLATAKRLGYISVTDVLKVV 790

RESULT 8
A35049
Ankyrin 1, erythrocyte splice form 2 - human
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
C:Species: Homo sapiens (man)
C:Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C:Accession: A35049
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370; PMID:1689849
A:Accession: A35049
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1880 <LAM>
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing; cytoskeleton
F:2-1513,1676-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MA2>
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>

```

F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match 14.6%; Score 844; DB 2; Length 1880;
Best Local Similarity 30.1%; Pred. No. 2.5e-44;
Matches 242; Conservative 121; Mismatches 263; Indels 177; Gaps 21;

QY 3 GRKDVVEYLQNGASVQARDGGGLIPLHNACSFQHAEEVYVNLRLRHGADPNARDNNWYTPL 62
DB 89 GQDEVVRELNVNGANVNAQSGKFTPLVMAQENHLEVVKFLLENGANQNVATEDGETPL 148
QY 63 HEAAIKGKIDVICVILQHCAPETI-----RNTDORTALDLADPSAKAVLTCEYKKD 113
DB 149 AVAALQOGENVVAHLINTGKGVRLPALHIAARNDDRTA-----AVLLQNDPNP 199
QY 114 ELLESARSNEEKMMALLTPLNVNCHASD-----GRKS-----TPLHLAAGY 155
DB 200 DVL--SKTG-----FTPLHIAHYENLVNAQLLNLRGSSVNFPTQNGITPLHIASRR 249
QY 156 NRKIVQLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKGACVNAWMLQFTPL 215
DB 250 GNVMVRLLDGQAETKTDELTPHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 309
QY 216 HEAAKNRVEVCSLLSYGADPTLLNCHNKSATDLAPTQLKERLAYEFKGHSLLQAARE 275
DB 310 HMAQGDHLQDVLQDYDE-----ID----- 332
QY 276 ADVTRIKKHLSEMVNFKHPOTHETALHCAASAPYKPKQICELLRKGANINEXTKFEFL 335
DB 333 -DIT-----LDHLTPLVHAA---HCGHHRVAKVLLDKGAKPNSRALNGF 372
QY 336 TPLHVAASEKAHNDVVEVYVVKHAKYNALDNLGQTSILHRAAYCGHLQTCRLLSYCDPNI 395
DB 373 TPLHIAACKNHVRVMBELLTKTASIDAVTESGLTPLHVASFMGHLPVIVKNLQIQRGASPNV 432
QY 396 ISLQGFALQMG-----NENVQOLIQE-----GTSLSGNSADRLQLE- 432
DB 433 SNVKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPPLHCAARIGHTNMVKLLLEN 492
QY 433 -----AAKAGDVETVKKLCTVQ--SVNCRDIEGROSTPLHFAAGYNRVVS 474
DB 493 NANPNLATTAGHTPLHIAAREGHVETVLALLEKEASQACMTKKG--FTPLHVAAYKGYR 550
QY 475 VVEYLLQHGADVIAKDKGLVPLHNACSYGHYEVTELLVKGACVNAWMLQFTPLHFA 533
DB 551 VAELLERDAHPNAAGKNGLTPLHVAHVHNNLDIVKLLPLRGSGSPH--SPAWNCGYTPHIA 609
QY 534 AAKGKYEICLLQHGADPTPKNRDNGTPLDL--VKDGDITDIQDLRGDAALLDAKKGCL 592
DB 610 AKQNVETVARSLLQYGGSSANESVQGVTPPLHIAAQEGHAEWALL----- 654
QY 593 ARVKKLSSPDNVNCRDPTQGRHSRSTPLHIAAGYNLEVAEYLLQHGADVNAQDKGGLTPLHN 652
DB 655 -----LSKQANGNLNKGSG--LTPLHLVAQEGHPVADVILKHHGVMDVATTRMGYTPH 707
QY 653 AASGYHVDVAALLIKYNACYNATDKWAFPLHEAAQKGRTOLCALLAHGADPTLKNQEG 712
DB 708 ASHYGNIKLVKFLQHQADVNAKTKLGYSPHQAAQOQGHDIPTVTLKLLKNGASPNVSSDG 767

QY 713 QTPLDL-----VSADDSALLT 729
DB 768 TTPLAIKRLGYSIVTDVLKVV 790

RESULT 9
SJHUK

N:Ankyrin 1, erythrocyte splice form 1 - human
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
C:Accession: S08275; A33219; PC2220; A35443
R:Lux, S.E.; John, K.M.; Bennett, V.
Nature 344, 36-42, 1990
A:Title: Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure
A:Reference number: S08275; MUID:90158830; PMID:2137557
A:Accession: S08275
A:Molecule type: mRNA
A:Residues: 1-1881 <LUI>
A:Cross-references: EMBL:X16609; NID:g28701; PIDN:CAA34610.1; PID:g28702
A:Accession: A33219
A:Molecule type: protein
A:Residues: 2-7, 'X', '9-17', 'X', '19-20', 'T', '22-30', '733-749', 'A', '751-753', '828-833', 'X', '835-855', 'X', '1367', '1383-1427', '1601-1630', '1686-1698', 'P', '1700', '1763-1772' <LUX>
A:Note: 845-Arg and 1392-Thr were also found
R:Hermann, J.; Barel, M.; Frade, R.
Biochem. Biophys. Res. Commun. 204, 453-460, 1994
A:Title: Human erythrocyte ankyrin, a cytoskeleton component, generates the p57 membr
A:Reference number: PC2220; MUID:95071348; PMID:7526850
A:Accession: PC2220
A:Molecule type: protein
A:Residues: 910-929 <HER>
R:David, L.H.; Bennett, V.
J. Biol. Chem. 265, 10589-10596, 1990
A:Title: Mapping the binding sites of human erythrocyte ankyrin for the anion exchang
A:Reference number: A35443; MUID:90285190; PMID:2141335
A:Accession: A35443
A:Molecule type: protein
A:Residues: 'X', '5', 'X', '7-12', '403-417', 'X', '419-422', 'H', '424', 'LQ'; '797-800', 'L', '802-814', '862-8
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing; phosphoprotein
F:2-1881/Product: ankyrin 1, erythrocyte form 1 #status predicted <MAT1>
F:2-1512,1675-1881/Product: ankyrin 2, erythrocyte form 2 #status predicted <MAT2>
F:2-827/Domain: 89K #status predicted <DOM1>
F:2-827/Region: anion exchange protein binding
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>
F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

F:828-1382/Domain: 62K #status predicted <DOM2>
F:828-1382/Region: spectrin binding
F:1383-1881/Domain: 55K #status predicted <DOM3>

Query Match 14.6%; Score 842; DB 1; Length 1881;
Best Local Similarity 29.9%; Pred. No. 3.4e-44;
Matches 240; Conservative 122; Mismatches 264; Indels 177; Gaps 20;

```
QY 3 GRKDVVEYLQNGASVOARDGGILPLHNACSFHAEVNVNLLRHGADPNARDNNWYTP 62
Db 89 GODEVRELNYGAVNAQSQKGTPLMAAQNHLVVKFLENGANQNVATEDGFTPL 148
QY 63 HEAAIKGIDVICVLLQHGAEPTI-----RNTDGRALDADPSAKAVLTGEYK 113
Db 149 AVALQOQHENVVAHLINYGTKGVRLPALHIAARNDDRTA-----AVLQNDPNP 199
QY 114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS-----TPHLAAGY 155
Db 200 DVL--SKTG-----FTPLHIAHYENLVNVAQLLLNRGASVNFTPONGITPLHIA 249
QY 156 NRKIVQLLQHGADVHAKDKGDLVPLHNACSYGHEVTELLVKGACVNAQMDLWQFTPL 215
Db 250 GNVIMVRLLLDRGAQIETKTKDELTPHCAARNGHVRISSEILLDHGAPIQAKTKNGLSPI 309
QY 216 HEAAKRNVEVCSLLSYGADPTLLNCHNKSADIDLAPTQPKERLAYEFKSHSLQAARE 275
Db 310 HMAAQGDHLDCVRLLOQYNAE-----ID----- 332
QY 276 ADVTRIKKHLISLEMVNFKHPQTHETALHCAASPYPKRKOICELLRLKGANINEKTEFL 335
Db 333 -DIT-----LDHLTPLHVA-----HCGHHRVAKVLLDKGAKPNRSLNGF 372
QY 336 TPLHVAASEKAHNDVVEVVKHKAQVNDLQGTSLHRAAYCGHLQTCRLLSYGCDDPNI 395
Db 373 TPLHIAKKNHVRVMEILLTKTGASIDAVTESGLTPLHVAQEGHVPADVLLKHGVVDAITRNGYPLHV 432
QY 396 ISLQGTALOM-----NENVOQLLOE-----GISLGNSEADRLLE- 432
Db 433 SNVKVETPLHMAARAGHTEVAKVLLQNKAKVNAKAKDDQTPHCAARIGHTNMVKLLLEN 492
QY 433 -----AAKAGDVEIVKKLCTVQ--SVNCRDIEGRQSTPLHFAAGYNRVS 474
Db 493 NANPNLATTAGHTPLHIAAREGHVEITVALLLEKESQACMTKG--FTPLHVAAKYGVKVR 550
QY 475 VVEYLQHGADVIAKDKGGLVPLHNACSYGHEVVAELVVKHGAQVNVADLWK--FTPLHEA 533
Db 551 VAELLERDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLPRGSGPH--SPAWNGYTPHIA 609
QY 534 AAKGKVEICKLLQHGADPTKKNRDGNTPDL-LVKDGTDIQDLRGDAALLDAAKGCL 592
Db 610 AKQOQVEVARSLLQYSGSANAESVQGVTPHLAAQEGHAEVWALL----- 654
QY 593 ARVKKLSSPDNNCRDQGRHSTPLHLAGYNNLEVAEYLLQHGADVNAQDKGLIPLHN 652
Db 655 -----LSKQANGNLNKG--LTPHLVQAEGHVPADVLLKHGVVDAITRNGYPLHV 707
QY 653 AASYGHVDVAALLIKYNACVNATDKWAFPTPLHFAAQGRTOICALLLHGAADPTLNQEG 712
Db 708 ASHYGNIKLVKFLQHOADVNAKTILGYSPFHQAQOQGHDTIVTLLKNGASPNVSSDG 767
QY 713 QTPDL-----VSADDVSALLT 729
Db 768 TPLAIAKRLIGYISVDLKVVT 790
```

RESULT 10

S37771

ankyrin, erythrocyte - mouse

C:Species: Mus musculus (house mouse)

C:Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999

C:Accession: S37771

R:Birkenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.

J. Biol. Chem. 268, 9533-9540, 1993

A:Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found
A:Reference number: S37771; MUID:93252825; PMID:8486643

A:Accession: S37771

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1848 <BIR>

A:Cross-references: EMBL:X69063; NID:g311816; PIDN:CAA48801.1; PID:g311817

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

F:48-80/Domain: ankyrin repeat homology <AN01>

F:81-113/Domain: ankyrin repeat homology <AN02>

F:114-146/Domain: ankyrin repeat homology <AN03>

F:147-175/Domain: ankyrin repeat homology <AN04>

F:176-208/Domain: ankyrin repeat homology <AN05>

F:209-241/Domain: ankyrin repeat homology <AN06>

F:242-274/Domain: ankyrin repeat homology <AN07>

F:275-307/Domain: ankyrin repeat homology <AN08>

F:308-340/Domain: ankyrin repeat homology <AN09>

F:341-373/Domain: ankyrin repeat homology <AN10>

F:374-406/Domain: ankyrin repeat homology <AN11>

F:407-439/Domain: ankyrin repeat homology <AN12>

F:440-472/Domain: ankyrin repeat homology <AN13>

F:473-505/Domain: ankyrin repeat homology <AN14>

F:506-538/Domain: ankyrin repeat homology <AN15>

F:539-571/Domain: ankyrin repeat homology <AN16>

F:572-604/Domain: ankyrin repeat homology <AN17>

F:605-637/Domain: ankyrin repeat homology <AN18>

F:638-670/Domain: ankyrin repeat homology <AN19>

F:671-703/Domain: ankyrin repeat homology <AN20>

F:704-736/Domain: ankyrin repeat homology <AN21>

F:737-769/Domain: ankyrin repeat homology <AN22>

F:770-802/Domain: ankyrin repeat homology <AN23>

Query Match

14.6%; Score 839.5; DB 2; Length 1848;

Best Local Similarity 29.5%; Pred. No. 4.7e-44;

Matches 237; Conservative 127; Mismatches 262; Indels 177; Gaps 20;

```
QY 3 GRKDVVEYLQNGASVOARDGGILPLHNACSFHAEVNVNLLRHGADPNARDNNWYTP 62
Db 93 GODEVRELNYGAVNAQSQKGTPLMAAQNHLVVKFLENGANQNVATEDGFTPL 152
QY 63 HEAAIKGIDVICVLLQHGAEPTI-----RNTDGRALDADPSAKAVLTGEYK 113
Db 153 AVALQOQHENVVAHLINYGTKGVRLPALHIAARNDDRTA-----AVLQNDPNP 203
QY 114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS-----TPHLAAGY 155
Db 204 DVL--SKTG-----FTPLHIAHYENLVNVAQLLLNRGASVNFTPONGITPLHIA 253
QY 156 NRKIVQLLQHGADVHAKDKGDLVPLHNACSYGHEVTELLVKGACVNAQMDLWQFTPL 215
Db 254 GNVIMVRLLLDRGAQIETKTKDELTPHCAARNGHVRISSEILLDHGAPIQAKTKNGLSPI 313
QY 216 HEAAKRNVEVCSLLSYGADPTLLNCHNKSADIDLAPTQPKERLAYEFKSHSLQAARE 275
Db 314 HMAAQGDHLDCVRLLOQYNAE-----ID----- 336
QY 276 ADVTRIKKHLISLEMVNFKHPQTHETALHCAASPYPKRKOICELLRLKGANINEKTEFL 335
Db 337 -DIT-----LDHLTPLHVA-----HCGHHRVAKVLLDKGAKPNRSLNGF 376
QY 336 TPLHVAASEKAHNDVVEVVKHKAQVNDLQGTSLHRAAYCGHLQTCRLLSYGCDDPNI 395
Db 377 TPLHIAKKNHVRVMEILLTKTGASIDAVTESGLTPLHVAQEGHVPADVLLKHGVVDAITRNGYPLHV 436
QY 396 ISLQGTALOM-----NENVOQLLOEISLGNSEAD-----ROLLE- 432
Db 437 SNVKVETPLHMAARAGHTEVAKVLLQNKAKVNAKAKDDQTPHCAARIGHTNMVKLLLEN 496
QY 433 -----AAKAGDVEIVKKLCTVQ--SVNCRDIEGRQSTPLHFAAGYNRVS 474
Db 497 GASPNIATTAGHTPLHIAAREGHVDVALLLEKESQACMTKKG--FTPLHVAAKYGVKVR 554
```


Best Local Similarity 26.18; Pred. No. 2.1e-42;		Matches 301; Conservative 156; Mismatches 418; Indels 279; Gaps 36;	
Qy	3	GRKDVVEYLLQNGASVQARDGGGLPLHNAACSGHAEVNVNLLRHGADPNARDNNVTPL 62	
Db	77	GHSVEVRELIHQAOVDAATRGKNTALHIAAGSLIVTLVENGANVNVOSVNGFTPL 136	
Qy	63	HEAATKGRIDVCI VLLQHGAEPTIRNTDRTALDLPDSAKAVLTGKDKDELLESARS 122	
Db	137	YMAAQENHEEVVYLLKHGAQALSTEDGFTPLA-----LQQG 176	
Qy	123	NEEKMAALLPLNVNCHASDGRKSTP-LHLAAGYNRVKIVOLLQHGADVHAKDKGLVP 181	
Db	177	HDRVVAVILLE-----NDSKGRVRLPALHIAAKDDTTAATLLQNEHNPVDTSKSFT 230	
Qy	182	LHNACSYGHEYTELLVHKGCNVAMDLMQFTPLHEAASKRVEVCSLLLSYGADPTLLN 241	
Db	231	LHIAHYGHENVGOLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLSRGA---IID 287	
Qy	242	CHNKSAIDLAPTQLKERLAYEFKSHLSLOAREADVTRIKKHLISLEMVNPKHQTHETA 301	
Db	288	SRTKDLL-----TP 296	
Qy	302	LHCAASPYPKKQICELLRLKGANINEKTEFLPLHVASEKAHNDVVEVVKHEAKVN 361	
Db	297	LHCAARSGH---DQVVDLLVVOGAPISAKTKNGLAPLHMAAGDHDVDAARTLLYHRAVD 353	
Qy	362	--ALDNLGOTS LHRAAYCGHLOTCLRLLSYGCDDPNIIISLOGFTALOMGNEN----- 410	
Db	354	DVTVDYL--TPLHVAACHGVHVKVALLDSDRSPNSRALNGFTPLHIAACKNRKRWELL 411	
Qy	411	-----VQQLQEGIS-----LNSEADROLLEAKAGDVETVKKICTVQSVNCRDIE 457	
Db	412	LKYRAAIEATTESGLTPLHVAAFMGAINIVYLLQOGANPDVETV----- 456	
Qy	458	GROSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNAACSYGHEVAELLVKHGA 517	
Db	457	-RGETPLHIAAARANTQDVRVLRNGAKYDAQARELOTPHLHSLRGNITDIIVILLQAGA 515	
Qy	518	VNVADLWFTPLHEAAAKGYEICKLLOHGADPTKKNRDNTPDLV-KGDDTDIQL 576	
Db	516	NSNATTRDNYSPHLIAAECQEBEAVAGILLDHNAKDTLLPKKGTPLHLASKYGNLEVR 575	
Qy	577	L--RGDAALLDAACKGLARVKLSSPDVNVNCHDTQGRHS-TPLHIAAGYNNLEVAEYLL 633	
Db	576	LLERGTPV-----DIEGKNQVTPHLVAAHYNNDKVAMLL 610	
Qy	634	OHGADVNAQDKGGLPLHNAASYGHVDVAALLIKYNACVNAATDKWAFTPLHEAAQKGRTO 693	
Db	611	ENGASAKAAAKNGYTPHLHIAAKKNQMEIATLLQFRADPNAKSRAGFTPLHLSAQEGHKE 670	
Qy	694	LCALLAHGADPTLKNOEGTPLDLVSAD-----VSALLTAAMPSPALPS 739	
Db	671	ISGLLIENGSDVGAKANGLTAMHLCAQEDHPVPAQILYNNAGAEINSKTNAGYTPHLV-A 729	
Qy	740	CYKPOVLNGVR---SPGATADALSSGSPSSLSAASSLDNLNSGFSSELSSVSSSGTEG 796	
Db	730	CHFGQ-LNVKFLVENGADVGEKTRASYTPHQAAGOHNNCVRYLLE-----NC 778	
Qy	797	ASSLEKKEVPGVDSITQFVRNLG-----LEHLMDFIREQOITLDLVEMGHKELKEIGIN 852	
Db	779	ASPNEQTATGTPLSIAQ---RLGYSVVETLRTVTITVTITVTVDERYKQPNPEAMN 835	
Qy	853	-----AYGRHKLI-----KGERLISQOGLNPVLTNTSG 884	
Db	836	ETMFSEDEGAHAEVAAHAEKDFSDNLTOGLQDSTGVHMIHTGQLQRQSOELENGG 895	
Qy	885	-----STILIDLSPDKEFQSEEMQS-----TVREHRDG--GHAGGIFNRYNITLKIO 932	
Db	896	AIPKINS-----GMSP-EKEPAKTAPVATSPATNSQSFIAPIRAGSISGQFOQPLH 950	
Qy	933	KVCNKLWERYTHRRKE--VSEENHNHANERMLPHGS-----PFVNAIHHGF-----DE 980	
Db	951	GAGPEDNLELVRAQHPIINAGYNDNGVAMLENHADNVPIGHVHTQPSFLISFLVDA 1010	
Qy	981	RHAYICGMFCAGIYFAENSKSNQYVYIGGTCGCPVHKDRSCYICHROLLFCRVTLGKS 1040	
Db	1011	RGAMRGCRHSRGVRIIVPPRKASQ-----PIR-----VTCRY-LRKD 1046	
Qy	1041	FLOFSAMKMAHSP 1054	
Db	1047	-----KLAHPP 1053	
RESULT 14			
T42691			
hypothetical protein DkfZp434D2328.1 - human (fragment)			
C:Species: Homo sapiens (man)			
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000			
R:Accession: T42691			
R:Bloeker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.			
submitted to the Protein Sequence Database, November 1999			
A:Reference number: Z22230			
A:Accession: T42691			
A>Status: preliminary			
A:Molecule type: mRNA			
A:Residues: 1-791 <AAA>			
A:Cross-references: EMBL:AL133087			
A:Experimental source: adult testis; clone DkfZp434D2328			
C:Genetics:			
A:Note: DkfZp434D2328.1			
Query Match 10.9%; Score 628; DB 2: Length 791;			
Best Local Similarity 27.7%; Pred. No. 2.4e-31;			
Matches 232; Conservative 119; Mismatches 313; Indels 174; Gaps 28;			
Qy	3	GRKDVVEYLLQNGASVQARDGGGLPLHNAACSGHAEVNVNLLRHGADPNARDNNVTPL 62	
Db	14	GOINVYKHLNLGVEIDINVYGNLTALHACYNQDQAVVNELIDYGANVGNPNNGFTPL 73	
Qy	63	HEAATKGRIDVCI-VLLOHGAEPTIRNTDRTALDLPDSAKAVLTGKDKELL----- 116	
Db	74	HFAAATHGALCLLELVNAGDVNIQSKGKSPHLM-----TAVHGRFTRSTLIQNGG 127	
Qy	117	-----ESARSGNEEKMMALLT-----PLN---VNCHA----- 140	
Db	128	ETDCVDKDGNTPLHVAARYGHELLINTLTSGADTKACGIHSMFPLHIAALNAHSDCCR 187	
Qy	141	--SDGRK-----LHLAGYNRVKIVOLLQHGADV 172	
Db	188	LLSGQKYSIVLSFNEHVSAGFEIDTPDKFGRVCLHAAAAGNVCEIKLLQSSGADFH 247	
Qy	173	AKDKGLVPLHNAACSYGHEYTELLVHKGCNVAMDLMQFTPLHEAASKRVEVCSLLLS 232	
Db	248	KKDKCGRTPLHYAAANCHFHCTETLVTTGANVNETDDMGRTALHYAAASD----- 297	
Qy	233	YCADPTLL-NCHNKSAIDIAPTPOLKER---LAYEFKSHLSLOAREADVTRIKKHL 288	
Db	298	MDRNTKILIGNAHNDE-ELERARELKEATLCLEF---LLQ----- 335	
Qy	289	MYNKHPOTHE---TALHCAASPYPKKQICELLRLKGANINEKTEFLT-PLHVAS 342	
Db	336	--NDANPSIRDEKGYNSIHYAAAYGH---ROCLELLELERTNSGFESDSGATKSPHL 390	
Qy	343	EKAHNDVVEVVKHKAQVNVNLDNLGQTSUHLRAAYCGHLOTCLRLLSYGCDDPNII 402	
Db	391	YNGHQAQLEVLQLSDLDIRDEKGTALDLAFAFKGTECEVALIN-----QGAS 440	
Qy	403	ALOMGNENVQOLLOEGISGNSSEADROLLEAAKAGDVETVKKICTVQSVNCRDIEGR 462	
Db	441	IFVKNVTRTPLHVASVINGHTLCLLELIAD-----NPEADVVDKAGQ--T 487	
Qy	463	PLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNAACSYGHEVAELLVKHGA 522	
Db	488	PLMLAVAYGHIDAVSLLLEKEANVDTVDILGCTALHGTMTGHECVQMLLQEVSI 547	

Query Match	9.5%;	Score 545.5;	DB 2;	Length 1411;
Best Local Similarity	23.7%;	Pred. No. 9.2e-26;		
Matches 239;	Conservative 136;	Mismatches 378;	Indels 257;	Gaps 33;
OY	11	LLONGASVOARDGGIPLHNACSGFHAENVNLLR-----HGADPNARDNNWYTFPHE	64	
Db	486	LIONGANVSETFELGRCAITHAASAGNYDVGELLNKDINLLKADKN-----GYTFLHI	540	
OY	65	AATKGKIDVICVLLQHGAEPTIRNTDGRALTADLPDSAKAVLTGEYKKDBELLESARS	124	
Db	541	ADSNKNDNFVFMELIGNADNV-----RTKSDLEFTP-----LHLAARRDLT	581	
OY	125	EKMALLTPLNVNCHASDGRKSTPHLAAGYNRVKIVOLLQHGADVHAKDGDVPLPHN	184	
Db	582	DVTQTLDITEIDLNADQKSGTFLHLSSTSTSETAAILRTNVAVINIKSKVGLTPLHL	641	
OY	185	ACSYGHYEVELLVKIGACVNAWMDIQMFTPLHEAASKNRVEVGLLSYGADPTLLNCHN	244	
Db	642	ATLQNNLSYKLLAGKAGYLNDGDANGMTPHLYAAMTGNLEMDVFL-----N	689	
OY	245	KSAILDAPPPOLKE-----PLAYEFKGHSLQ-----	271	
Db	690	QOYININAAATKEKKWTPHLAILFKKNDVAERLLSDENINIRLETNGGINPLHASATGN	749	
OY	272	-----AAREADVTPI-KKHLISLEVMNFKPQTHETALHCAASPYPKRKQICELLRK	323	
Db	750	KOLVITELAKNADVTPLRLTSKGFS-----ALHLGIIG---KNEEIPFPLVEK	792	
OY	324	GANINIEKTEKFTPLPHRVASEKAHNDVVEVVVKEHAVNALDNLQTSLHRAAYCGHLQTC	383	
Db	793	GANVNDKNTSGVTPPLHFAAGLCKANIFRLLLSRGADIKAEIDINSMPIHEAVSNGHLETV	852	
OY	384	RLLLSYGCDPNIISLOGFTALQMGNEVNOQLLOE-GISLGNSEAFDQLLEAAKAGDVETV	442	
Db	853	RIILEK-----DPSLM-----NVKNIINREYPPFYLAVERKRYKDIFDYFVSKD----	893	

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 21:15:54 ; Search time 13.6325 Seconds
(without alignments)
3346.711 Million cell updates/sec

Title: US-09-843-159B-3
Perfect score: 5769
Sequence: 1 GFGKDVVEYLLONGASVQA.....AYPEYLITYQIMRPEGWDG 1100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5766	99.9	1166	1	TNK2_HUMAN
2	4889.5	84.8	1327	1	TNK1_HUMAN
3	875.5	15.2	4377	1	ANK3_HUMAN
4	860.5	14.9	3924	1	ANK2_HUMAN
5	842	14.6	1880	1	ANK1_HUMAN
6	837.5	14.5	1862	1	ANK1_MOUSE
7	707.5	12.3	1059	1	Y379_HUMAN
8	571	9.9	768	1	YB23_HUMAN
9	464.5	8.1	1401	1	LATA_LATMA
10	420.5	7.3	603	1	V162_FOWPV
11	394.5	6.8	747	1	V222_FOWPV
12	390.5	6.8	832	1	ANK3_HUMAN
13	380.5	6.6	668	1	V244_FOWPV
14	368	6.4	596	1	V024_FOWPV
15	353.5	6.1	1431	1	DAPK_HUMAN
16	351	6.1	542	1	V155_FOWPV
17	350	6.1	587	1	ASB2_HUMAN
18	334.5	5.8	775	1	ASB5_MOUSE
19	327.5	5.7	525	1	V228_FOWPV
20	322.5	5.6	776	1	ANK5_HUMAN
21	322	5.6	525	1	ASB3_MOUSE
22	321	5.6	436	1	V245_FOWPV
23	319	5.5	592	1	V246_FOWPV
24	310	5.4	528	1	PPGA_HUMAN
25	309	5.4	524	1	PPGA_MOUSE
26	309	5.4	568	1	PPGB_MOUSE
27	306.5	5.3	518	1	ASB3_HUMAN
28	305	5.3	735	1	RNSA_MOUSE
29	303.5	5.3	568	1	PPGB_BOVIN
30	301.5	5.2	567	1	PPGB_HUMAN
31	299.5	5.2	621	1	ANK6_HUMAN
32	291.5	5.1	950	1	ORP1_HUMAN
33	287.5	5.0	656	1	FEMI_CABEL

34	286.5	5.0	741	1	RNSA_HUMAN
35	285	4.9	461	1	V218_FOWPV
36	285	4.9	1083	1	Y112_YEAST
37	281.5	4.9	482	1	V232_FOWPV
38	281.5	4.9	898	1	KBF2_HUMAN
39	278	4.8	984	1	KBF1_CHICK
40	277.5	4.8	429	1	AS10_HUMAN
41	277	4.8	1964	1	NTC4_MOUSE
42	276.5	4.8	434	1	V219_FOWPV
43	275	4.8	347	1	GABC_MOUSE
44	274	4.7	382	1	GABB_MOUSE
45	272.5	4.7	968	1	KBF1_HUMAN

ALIGNMENTS

RESULT 1
TNK2_HUMAN STANDARD; PRT: 1166 AA.
ID TNK2_HUMAN
AC O9H2K2: O9HAS4: Q9H8F2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Tankyrase 2 (EC 2.4.2.30) (TANK2) (Tankyrase II) (TNKS-2) (TRP1-
interacting ankyrin-related ADP-ribose polymerase 2) (Tankyrase-like
protein) (Tankyrase-related protein).
GN TNKS2 OR TNKL OR TANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Fetal brain;
RC MEDLINE=21072550; PubMed=11205898;
RX Monz D., Munnia A., Comtesse N., Fischer U., Steudel W.-I., Feiden W.,
RA Glass B., Meese E.U.;
RT "Novel tankyrase-related gene detected with meningioma-specific
sera";
RL Clin. Cancer Res. 7:113-119(2001).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Breast carcinoma;
RC MEDLINE=21190090; PubMed=11294570;
RX Kulmov A.N., Kuprash D.V., Petrov V.N., Vdovichenko K.K.,
RA Scanlan M.J., Jongeneel C.V., Lagarkova M.A., Nedospasov S.A.;
RT "Cloning and characterization of TNKL, a member of tankyrase gene
family";
RL J. Biol. Chem. 276:17172-17180(2001).
[3]
RN SEQUENCE FROM N.A., AND SUBCELLULAR LOCALIZATION.
RP TISSUE=Liver;
RC MEDLINE=21264473; PubMed=11278563;
RX Lyons R.J., Deane R., Lynch D.K., Ye Z.-S.J., Sanderson G.M.,
RA Eyre H.J., Sutherland G.R., Daly R.J.;
RT "Identification of a novel human tankyrase through its interaction
with the adaptor protein Grb14";
RL J. Biol. Chem. 276:17172-17180(2001).
[4]
RN SEQUENCE FROM N.A., AND SUBCELLULAR LOCALIZATION.
RP TISSUE=Placenta;
RC MEDLINE=21443728; PubMed=11454873;
RX Kaminker P.G., Kim S.-H., Taylor R.D., Zebarjadian Y., Funk W.D.,
RA Morin G.B., Yaswen P., Campisi J.;
RT "TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes
rapid induction of cell death upon overexpression";
RL J. Biol. Chem. 276:35891-35899(2001).
[5]
RN SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH TRF1 AND
RP LNPEP/OTASE.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21661461; PubMed=11802774;

Q05823	homo sapien
Q9J517	fowlpox vir
P40480	saccharomyc
Q9J503	fowlpox vir
Q00653	homo sapien
Q04861	gallus gall
Q8WXI3	homo sapien
P31695	mus musculus
Q9J516	fowlpox vir
Q00421	mus musculus
Q00420	mus musculus
P19838	homo sapien

RA Sbodio J.I., Lodish H.F., Chi N.-W.;
 RT "Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1
 RT (telomere-repeat-binding factor 1) and TRAP (insulin-responsive
 RT aminopeptidase).";
 RL Biochem. J. 361:451-459(2002).
 RN [6]
 RN SEQUENCE FROM N.A.
 RA Yin Y., Gelmann E.P.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RN SEQUENCE FROM N.A.
 RA Tracey A.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RN SEQUENCE OF 838-1151 FROM N.A.
 RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Fogliya S., Komai F., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RN FUNCTION, AND ADP-RIBOSYLATION.
 RP MEDLINE-21602874; PubMed=11739745;
 RA Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
 RT "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2
 RT at human telomeres.";
 RL Mol. Cell. Biol. 22:332-342(2002).
 CC -!- FUNCTION: May regulate vesicle trafficking and modulate the
 CC subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP
 CC activity and can modify TRF1, and thereby contribute to the
 CC regulation of telomere length.
 CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboyl](N)-acceptor =
 CC nicotinamide + [ADP-D-riboyl](N+1)-acceptor
 CC -!- SUBUNIT: Oligomerizes and associates with TNKS. Interacts with the
 CC cytoplasmic domain of LNP/OTase in SLC2A4/GLUT4-vesicles. Binds
 CC to the N-terminus of Grb14 and TRF1 with its ankyrin repeat
 CC region.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
 CC with juxtanuclear SLC2A4/GLUT4-vesicles. Also found around the
 CC pericentriolar matrix of mitotic centrosomes. During interphase, a
 CC small fraction of TNKS2 is found in the nucleus, associated with
 CC TRF1.
 CC -!- TISSUE SPECIFICITY: Highly expressed in placenta, skeletal muscle,
 CC liver, brain, kidney, heart, thymus, spinal cord, lung, peripheral
 CC blood leukocytes, pancreas, lymph nodes, spleen, prostate, testis,
 CC ovary, small intestine, colon, mammary gland, breast and breast
 CC carcinoma, and in common-type meningioma. Highly expressed in
 CC fetal liver, heart and brain.
 CC -!- PTM: ADP-ribosylated (-auto).
 CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
 CC -!- SIMILARITY: CONTAINS 15 ANK REPEATS.
 CC -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF305081; AAG25674.1; ALT_INIT.
 DR EMBL; AF264912; AAG44694.1; -
 DR EMBL; AF329696; AAK13463.1; -
 DR EMBL; AF342982; AAK25811.1; -
 DR EMBL; AF309033; AAK82330.1; -
 DR EMBL; AF438201; AAL40795.1; -
 DR EMBL; AL359707; CAC78760.1; -
 DR EMBL; AK023746; BAB14665.1; ALT_INIT.
 DR HSP; Q00420; IAWC.

DR Genew; HGNC:15677; TNKS2.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; ank; 20.
 DR Pfam; PF00536; SAM; 1.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 15.
 DR SMART; SM00454; SAM; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 15.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PS50105; SAM_DOMAIN; 1.
 KW Nuclear protein; Repeat; ANK repeat; ADP-ribosylation.
 FT REPEAT 57 89
 FT REPEAT 90 122
 FT REPEAT 123 155
 FT REPEAT 210 242
 FT REPEAT 243 275
 FT REPEAT 276 308
 FT REPEAT 363 398
 FT REPEAT 399 431
 FT REPEAT 432 464
 FT REPEAT 525 557
 FT REPEAT 558 590
 FT REPEAT 591 623
 FT REPEAT 678 710
 FT REPEAT 711 743
 FT REPEAT 744 776
 FT DOMAIN 873 936
 FT DOMAIN 1023 1162
 FT FLNAC 331 337
 FT FLNAC 357 361
 FT FLNAC 966 966
 SQ SEQUENCE 1166 AA; 126917 MW; 4C8B3B8D97CE704 CRC64;
 Query Match 99.9%; Score 5766; DB 1; Length 1166;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GFGKRDVVEYLLONGASVQARDGGLIPLHNACSFHAEVYVNLRLRHGADPNARDNNYT 60
 DB 67 GFGKRDVVEYLLONGASVQARDGGLIPLHNACSFHAEVYVNLRLRHGADPNARDNNYT 126
 QY 61 PLHEAAIKGIDVCIVLLOHGAETIRNTDGTALDADPSAKAVLTGEYKDELLSAR 120
 DB 127 PLHEAAIKGIDVCIVLLOHGAETIRNTDGTALDADPSAKAVLTGEYKDELLSAR 186
 QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLV 180
 DB 187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLV 246
 QY 181 PLHNACSYGHVEYTELLVKGACVNMADLWQFTPLHFAASKNRVVECSLLSYGADPTLL 240
 DB 247 PLHNACSYGHVEYTELLVKGACVNMADLWQFTPLHFAASKNRVVECSLLSYGADPTLL 306
 QY 241 NCHNKAIDLAPLPOLKERLAYEFKHSLLQAAAREADVTIRKKHLSLEMYNFKHPQTHET 300
 DB 307 NCHNKAIDLAPLPOLKERLAYEFKHSLLQAAAREADVTIRKKHLSLEMYNFKHPQTHET 366
 QY 301 ALHCAASPYPKRKQICEILLRKGANINEKTEFLTPHVASEKAHNDVVEVVKHEAKV 360
 DB 367 ALHCAASPYPKRKQICEILLRKGANINEKTEFLTPHVASEKAHNDVVEVVKHEAKV 426
 QY 361 NALDNLGQTSLHRAAYCGHLOTCRLLLSYGCDPNIIISLQGTALQMGNNVQQLLEGIS 420
 DB 427 NALDNLGQTSLHRAAYCGHLOTCRLLLSYGCDPNIIISLQGTALQMGNNVQQLLEGIS 486
 QY 421 LGNSEADROLLEAAKAGDVETVKKLTQVSVCNCRDIEGQSTPLHFAAGYNRVSVVEYLL 480
 DB 487 LGNSEADROLLEAAKAGDVETVKKLTQVSVCNCRDIEGQSTPLHFAAGYNRVSVVEYLL 546
 QY 481 QHGADVHAKDKGLVPLHNACSYGHVEYAEVLVKGACVNVADLWKFTEPLHEAAKGYE 540

Db 547 QHCADVHAKDGCLVPLHNAACSYGHVEAELLVKGAVVNVADLWKFPLHEAAAKGYE 606

Qy 541 ICKLLLOHGAADPTKKNRGNTPLDLVKDGDITDIQLLRCDALLDAAKGGLARVKKLSS 600

Db 607 ICKLLLOHGAADPTKKNRGNTPLDLVKDGDITDIQLLRCDALLDAAKGGLARVKKLSS 666

Qy 601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLHOGADVNAQDKGLPLHNAASYGHVD 660

Db 667 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLHOGADVNAQDKGLPLHNAASYGHVD 726

Qy 661 VAALLIKYNACVNAATDKWAFPLHEAOKGRQTLQALLAHGADPTLKNQEGQTPLDLVS 720

Db 727 VAALLIKYNACVNAATDKWAFPLHEAOKGRQTLQALLAHGADPTLKNQEGQTPLDLVS 786

Qy 721 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLDNLGS 780

Db 787 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLDNLGS 846

Qy 781 SFSELSVVSSSGTEGASSLEKKEVPGVDFSTQFVRNGLGLEHMDIPEREQITLDVLVE 840

Db 847 SFSELSVVSSSGTEGASSLEKKEVPGVDFSTQFVRNGLGLEHMDIPEREQITLDVLVE 906

Qy 841 MGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Db 907 MGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPYLTLNTSGSGTILIDLSPDDKEFQ 966

Qy 901 SYVEEMOSTVREHROGGHAGGIFNRYNLIKIOKCNKWLERYTHRRKEVSEENHNHANE 960

Db 967 SYVEEMOSTVREHROGGHAGGIFNRYNLIKIOKCNKWLERYTHRRKEVSEENHNHANE 1026

Qy 961 RMLFHGSPFVNAILHKGDFDERHAYIGGMFAGIYFAENSCKSNQYVYGGTGCPVHKD 1020

Db 1027 RMLFHGSPFVNAILHKGDFDERHAYIGGMFAGIYFAENSCKSNQYVYGGTGCPVHKD 1086

Qy 1021 RSCYCHROLLFCRYTLGKSFLOFSAMKMAHSPGHSHSVTGRPSVNGLALAEYVYIRGQ 1080

Db 1087 RSCYCHROLLFCRYTLGKSFLOFSAMKMAHSPGHSHSVTGRPSVNGLALAEYVYIRGQ 1146

Qy 1081 AYPEYLITYQIMRPEGMDVG 1100

Db 1147 AYPEYLITYQIMRPEGMDVG 1166

RESULT 2

TKN1 HUMAN

ID TKN1 HUMAN STANDARD; PRT: 1327 AA.

AC 095271; 095272;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE Tankyrase 1 (EC 2.4.2.30) (TANK1) (Tankyrase 1) (TNKS-1) (TRF1-interacting ankyrin-related ADP-ribose polymerase).

GN TNKS OR TNKS1 OR TINI OR TIN1 OR PARPL.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).

RC TISSUE=Testis;

RX MEDLINE=99040105; PubMed=9822378;

RA Smith S., Giriat L., Schmitt A., de Lange T.;

RT "Tankyrase, a poly(ADP-ribose) polymerase at human telomeres.";

RL Science 282:1484-1487(1998).

RN [2]

RP SUBCELLULAR LOCALIZATION.

RX MEDLINE=99454782; PubMed=10523501;

RA Smith S., de Lange T.;

RT "Cell cycle dependent localization of the telomeric PARP, tankyrase, to nuclear pore complexes and centrosomes.";

RL J. Cell Sci. 112:3649-3656(1999).

RN [3]

RP FUNCTION, AND PHOSPHORYLATION.

TX MEDLINE=20556282; PubMed=10988299;

Chi N.-W., Lodish H.F.;

"Tankyrase is a golgi-associated mitogen-activated protein kinase substrate that interacts with IRAP in GLUT4 vesicles.";

J. Biol. Chem. 275:38437-38444 (2000).

[4]

FUNCTION, AND MUTAGENESIS OF HIS-1184 AND GLU-1291.

MEDLINE=21602874; PubMed=11739745;

Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;

"Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2 at human telomeres.";

Mol. Cell. Biol. 22:332-342(2002).

CC FUNCTION: May regulate vesicle trafficking and modulate the subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP activity and can modify TRF1, and thereby contribute to the regulation of telomere length.

CC CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboseyl](N)-acceptor = nicotinamide + [ADP-D-riboseyl](N+1)-acceptor.

CC SUBUNIT: Oligomerizes and associates with TNKS2. Interacts with the cytoplasmic domain of INPEP/Otase in SLC2A4/GLUT4-vesicles. Binds to the N-terminus of telomeric TRF1 via the ANK repeats.

CC SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and with juxtanuclear SLC2A4/GLUT4-vesicles. A minor proportion is also found at nuclear pore complexes and around the pericentriolar matrix of mitotic centrosomes. During interphase, a small fraction of TNKS is found in the nucleus, associated with TRF1.

CC ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be produced by alternative splicing.

CC TISSUE SPECIFICITY: Ubiquitous; highest levels in testis.

CC PTM: Upon insulin-stimulation, phosphorylated on serine residues by MAPK kinases.

CC PTM: ADP-ribosylated (-auto).

CC SIMILARITY: BELONGS TO THE PARP FAMILY.

CC SIMILARITY: CONTAINS 15 ANK REPEATS.

CC SIMILARITY: CONTAINS 1 SAM DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AF082556; AAC79841.1; -

EMBL: AF082557; AAC79842.1; -

EMBL: AF082558; AAC79843.1; -

EMBL: AF082559; AAC79844.1; -

HSSP: Q00420; IAWC.

Genew: HGNC:11941; TNKS.

MIM: 603303; -

InterPro: IPR002110; ANK.

InterPro: IPR001660; SAM.

Pfam: PF00423; ank; 21.

Pfam: PF00536; SAM; 1.

PRINTS: PR01415; ANKYRIN.

SMART: SM00248; ANK; 15.

SMART: SM00454; SAM; 1.

PROSITE: PS50088; ANK_REPEAT; 15.

PROSITE: PS50297; ANK_REPEAT_REGION; 1.

PROSITE: PS50105; SAM_DOMAIN; 1.

KW Transferase; Glycosyltransferase; NAD; Golgi stack; Telomere; Nuclear protein; Repeat; ANK repeat; ADP-ribosylation; Phosphorylation; Alternative splicing.

KW REPEAT 215 247 ANK 1.

FT REPEAT 248 280 ANK 2.

FT REPEAT 281 313 ANK 3.

FT REPEAT 368 400 ANK 4.

FT REPEAT 401 433 ANK 5.

FT REPEAT 434 466 ANK 6.

FT REPEAT 521 556 ANK 7.

FT REPEAT 557 589 ANK 8.

FT REPEAT 590 622 ANK 9.

```
FT REPEAT 683 715 ANK 10.
FT REPEAT 716 748 ANK 11.
FT REPEAT 749 781 ANK 12.
FT REPEAT 836 868 ANK 13.
FT REPEAT 869 901 ANK 14.
FT REPEAT 902 934 ANK 15.
FT DOMAIN 1030 1089 SAM.
FT DOMAIN 1176 1327 PARP.
FT DOMAIN 9 14 POLY-HIS.
FT DOMAIN 27 34 POLY-PRO.
FT DOMAIN 128 134 POLY-SER.
FT DOMAIN 137 145 POLY-SER.
FT VARSPLIC 641 643 EST -> GHS (IN ISOFORM 2).
FT VARSPLIC 644 1327 MISSING (IN ISOFORM 2).
FT MUTAGEN 1184 1184 H->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
WITH A-1291.
FT MUTAGEN 1291 1291 E->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
WITH A-1184.
FT SEQUENCE 1327 AA: 142010 MW: 814DE985C710B957 CRC64;
Query Match 84.8%; Score 4889.5; DB 1; Length 1327;
Best Local Similarity 83.0%; Pred. No. 3-2e-292;
Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

QY 1 GFGKDVVEYLLQNGASVQARDGGGLIPHNACSFHAEVYVNLRLRHGADPNARDNNYT 60
DB 225 GFGKDVVEYLLQNGASVQARDGGGLIPHNACSFHAEVYVNLRLRHGADPNARDNNYT 284
QY 61 PLHEAAIKGKIDVICVILLOHGAPEPTIRNTDGTALDLDAPSAAVLTGEYKDDLEESAR 120
DB 285 PLHEAAIKGKIDVICVILLOHGAPEPTIRNTDGTALDLDAPSAAVLTGEYKDDLEESAR 344
QY 121 SGNEEKMMALLTPLNYNCHASDGRKSTPLHLAAGYNRVKIVQLLOHGAADVHAKDGLV 180
DB 345 SGNEEKMMALLTPLNYNCHASDGRKSTPLHLAAGYNRVKIVQLLOHGAADVHAKDGLV 404
QY 181 PLHNACSYGHEYTELLVYKHGACVNDMLWQFTPLHEAASKNRVEVCVSLLSYGADPTLL 240
DB 405 PLHNACSYGHEYTELLVYKHGACVNDMLWQFTPLHEAASKNRVEVCVSLLSYGADPTLL 464
QY 241 NCHNSAIDLAPTPOLKERLAYEFKGHSLLOAAREADVTIRIKKHLSELMVNFKHPOTHET 300
DB 465 NCHGKSAMDAPTPOLKERLAYEFKGHSLLOAAREADVTIRIKKHLSELMVNFKHPOTHET 524
QY 301 ALHCAASAPYKPKQTCCELLLRKANINEKTEFLPLHVASEKAHNDVYVYVVEVVEAV 360
DB 525 ALHCAVASLHPKPKQTCCELLLRKANINEKTEFLPLHVASEKAHNDVYVYVVEVVEAV 584
QY 361 NALDNLGQTSLHRAAYCGHLCQTRALLSYGCDPNITSLQGTALONGENVQQLQEGIS 420
DB 585 NALDNLGQTSLHRAAYCGHLCQTRALLSYGCDPNITSLQGTALONGENVQQLQEGIS 644
QY 421 LGNSEADROLLEAAKAGDVETVKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
DB 645 IRTSDVDYRLLEASKAGDLETKQLCSSQNVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 704
QY 481 QHGADVHAKDKGLVPLHNACSYGHEYTELLVYKHGACVNDMLWQFTPLHEAASKNRVE 540
DB 705 QHGADVHAKDKGLVPLHNACSYGHEYTELLVYKHGACVNDMLWQFTPLHEAASKNRVE 764
QY 541 ICKLLLOHGAADPTKKNRCDGNTPLDLVKDGDTDIQLDGLDAAIDAAKKGCLARVKLSS 600
DB 765 ICKLLLOHGAADPTKKNRCDGNTPLDLVKDGDTDIQLDGLDAAIDAAKKGCLARVKLSS 824
QY 601 PNVNCRDQGRHSTPLHLAAGYNRLLEVAEYLLQHGADYNAQDKGLIPLHNAASYGHVD 660
DB 825 PNVNCRDQGRHSTPLHLAAGYNRLLEVAEYLLQHGADYNAQDKGLIPLHNAASYGHVD 884
QY 661 VAALLIKYNACVNDTKWAFPTPLHEAAGKRTQQLCALLAHGADPTLNQEGTPTLDIVS 720
DB 885 VAALLIKYNACVNDTKWAFPTPLHEAAGKRTQQLCALLAHGADPTLNQEGTPTLDIVS 944
QY 721 ADDVSALLTAAMPSPALPSCYKQVINGVRSPGATADALSSGSPSSLSAASSLDNLSSG 780
```

```
DB 945 ADDIRALLIDAMPPEALPTCFKQAT-----VVSASLISPASTPSCLSAASIDNLITG 997
QY 781 SFSELSVWSSSGTEGASSLEKK--EVPGVDFSTQVFNRLGLEHLMDFEREQITLDVL 838
DB 998 PLAEAVGGSNAGDGAAGTERKEGEVAGLDMNISQPLKSLGLEHLRDFETEQITLDVL 1057
QY 839 VEMGHKELKEIGINAYCHRUHLKIKGVERLISGOGLNPYLTLNTISGCTILIDLSPPDKE 898
DB 1058 ADMGHEELKEIGINAYCHRUHLKIKGVERLISGOGLNPYLTLNTISGCTILIDLSPPDKE 1117
QY 899 FQSVSEEMQSTVREHRDGGHAGIFNRYNLIKQKVNKKIKWERYTHRRKVESEENHHA 958
DB 1118 YQSVSEEMQSTVREHRDGGHAGIFNRYNLIKQKVNKKIKWERYTHRRKVESEENHHA 1177
QY 959 NERMLFHGSPFNVAIHKGFDERHAYIGMGFAGIYFAENSCKSNQYVYGGGTGCPVH 1018
DB 1178 NERMLFHGSPFNVAIHKGFDERHAYIGMGFAGIYFAENSCKSNQYVYGGGTGCPVH 1237
QY 1019 KDRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPGHHSVTPGRPSVNGLALEYVIYRG 1078
DB 1238 KDRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPGHHSVTPGRPSVNGLALEYVIYRG 1297
QY 1079 EQAYPEYLITYQIMRPE 1095
DB 1298 EQAYPEYLITYQIMRPE 1314
RESULT 3
ANK3_HUMAN STANDARD; PRT: 4377 AA.
AC Q12955;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 3 (ANK-3) (Ankyrin G).
GN ANK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain stem;
RX MEDLINE=95138209; PubMed=7836469;
RA Kordeli E., Lambert S., Bennett V.;
RT "AnkyrinG. A new ankyrin gene with neural-specific isoforms localized
at the axonal initial segment and node of Ranvier.";
RL J. Biol. Chem. 270:2352-2359(1995).
CC -!- FUNCTION: Membrane-cytoskeleton linker.
CC -!- ALTERNATIVE PRODUCTS: A number of isoforms are produced by
alternative splicing.
CC -!- TISSUE SPECIFICITY: Expressed in brain and other tissues.
CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U13616; AAA64834.1; -.
DR HSP: P55273; 1B78.
DR Genew: HGNC:494; ANK3.
DR MIM: 600465; -.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000488; Death.
DR InterPro: IPR000906; ZU5.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00531; death; 1.
```

```

DR Pfam; PF00791; ZUS; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS00088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat.
FT REPEAT 73 102 ANK 1.
FT REPEAT 106 135 ANK 2.
FT REPEAT 139 168 ANK 3.
FT REPEAT 172 201 ANK 4.
FT REPEAT 203 230 ANK 5.
FT REPEAT 234 263 ANK 6.
FT REPEAT 267 296 ANK 7.
FT REPEAT 300 329 ANK 8.
FT REPEAT 333 362 ANK 9.
FT REPEAT 366 395 ANK 10.
FT REPEAT 399 428 ANK 11.
FT REPEAT 432 461 ANK 12.
FT REPEAT 465 494 ANK 13.
FT REPEAT 498 527 ANK 14.
FT REPEAT 531 560 ANK 15.
FT REPEAT 564 593 ANK 16.
FT REPEAT 597 626 ANK 17.
FT REPEAT 630 659 ANK 18.
FT REPEAT 663 692 ANK 19.
FT REPEAT 696 725 ANK 20.
FT REPEAT 729 758 ANK 21.
FT REPEAT 762 791 ANK 22.
FT REPEAT 795 825 ANK 23.
FT DOMAIN 1519 1898 SER-RICH.
FT DOMAIN 4090 4174 DEATH.
SQ SEQUENCE 4377 AA; 480399 MW; F42379E55768B684 CRC64;

Query Match 15.2%; Score 875.5; DB 1; Length 4377;
Best Local Similarity 30.5%; Pred. No. 4.4e-45;
Matches 273; Conservative 126; Mismatches 362; Indels 135; Gaps 21;

QY 3 GRKDVVEYLLQNGASVQARDGGGLIPLINACSFHAEVNVLLLRHGADPNARDNWTPL 62
DB 85 GHVEVSELLQREANVDAAATKGNLTALHSLAGAEVVKVLYVTNGANVNAQSQNGETPL 144
QY 63 HEAAIKGIDIVLLQHGAEPTIRNTDGRATDLA-----DPSAKAVLGEYK---KDE 114
DB 145 YMAAQEHLEVVFKLLDNGASQSLEATDGTPLAVALQQHDQVVSLLLENDTKGKVLRLP 204
QY 115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAK 174
DB 205 ALHIAARKDDTKAAALLQNDNNADVESKSGFTPLHIAAHYGNINVTALLNRAAAVDFT 264
QY 175 DKGLVPLHNACSGHYEVTELLVKGACVNMDLWQFTPLHGAASKNRVEVCSLLLSYG 234
DB 265 ARNDITPLHVASKRGANVMVKLLDRGAKIDAKTRDGLTPLHCGARSQHEQGVEMLLDRA 324
QY 235 ADPTLLNCHNKSAIDAPTLPOLKERLAYEFKGSLLQAAAREADVTTRIKKHLSEWNEFKH 294
DB 325 A-PLLSKTKN-----GLSPLHATQGD-----HLNCVQLLQHQ 356
QY 295 -----POTHETALHCAAAAPYKPKQICELLRLKGANINEKTKFELTPLHVAASEKAHN 347
DB 357 NVPVDDTNDYLTALHVAACHGHYK---VAKVLLDKKANPAKALNGFTPLHIACKKNRI 413
QY 348 DVVEVVKHAKVNALDNLGTSLHRAAYCHLOTCRLLLSYCGDNPDIISLQGTALQMG 407
DB 414 KVMELLKHGASIQAVTESGLTPTTHVAFMGHVIVISQLMHHGASPNNTNVRGETALHMA 473
QY 408 -----NENVOQLLOEGISLGNSEADRO--LLEAAKAGDVETVKKLCIV-QSVNCRDIEGR 459
DB 474 ARSQAEVVRVLVQDGAQVENAKAKDDTPLHISARLKGADIVQQLLQGGASPNATTSYG- 532
QY 460 QSTPLHFAAGYNRSVVVEYLLQHGADVHAKDKGGLVPLHNACSYGHVEVALLVKHGAVV 519

```


CC -1- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL
 CC CELLS THROUGHOUT THE BRAIN.
 CC -1- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES
 CC AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
 CC AND FUNCTION (POTENTIAL).
 CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X56957; CAA40278.1; -;
 CC EMBL; X56958; CAA40279.2; -;
 CC EMBL; Z26634; CAB42644.1; -;
 CC EMBL; M37123; AAG62828.1; -;
 CC PIR; S14533; S14533.
 CC PIR; A39643; A39643.
 CC PIR; B39643; B39643.
 CC PIR; S14569; S14569.
 CC HSSP; P42771; IDC2.
 CC Genew; HGNC:493; ANK2.
 CC MIM; 106410; -;
 CC InterPro; IPR002110; ANK.
 CC InterPro; IPR000488; Death.
 CC InterPro; IPR000906; ZU5.
 CC Pfam; PF00023; ank; 24.
 CC Pfam; PF00531; death; 1.
 CC Pfam; PF00791; ZU5; 1.
 CC PRINTS; PR01415; ANKYRIN.
 CC SMART; SM00248; ANK; 21.
 CC SMART; SM00005; DEATH; 1.
 CC SMART; SM00218; ZU5; 1.
 CC PROSITE; PS50088; ANK_REPEAT; 20.
 CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 CC PROSITE; PS50017; DEATH_DOMAIN; 1.
 KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
 KW Phosphorylation.
 FT REPEAT 63 92 ANK 1.
 FT REPEAT 96 125 ANK 2.
 FT REPEAT 129 158 ANK 3.
 FT REPEAT 162 191 ANK 4.
 FT REPEAT 193 220 ANK 5.
 FT REPEAT 232 261 ANK 6.
 FT REPEAT 265 294 ANK 7.
 FT REPEAT 298 327 ANK 8.
 FT REPEAT 331 360 ANK 9.
 FT REPEAT 364 393 ANK 10.
 FT REPEAT 397 426 ANK 11.
 FT REPEAT 430 459 ANK 12.
 FT REPEAT 463 492 ANK 13.
 FT REPEAT 496 525 ANK 14.
 FT REPEAT 529 558 ANK 15.
 FT REPEAT 562 591 ANK 16.
 FT REPEAT 595 624 ANK 17.
 FT REPEAT 628 657 ANK 18.
 FT REPEAT 661 690 ANK 19.
 FT REPEAT 694 723 ANK 20.
 FT REPEAT 727 756 ANK 21.
 FT REPEAT 760 789 ANK 22.
 FT REPEAT 793 822 ANK 23.
 FT DOMAIN 1773 1950 REPEAT-RICH REGION.
 FT REPEAT 1773 1784 REPEAT A.
 FT REPEAT 1785 1796 REPEAT A.
 FT REPEAT 1797 1808 REPEAT A.
 FT REPEAT 1809 1820 REPEAT A.
 FT REPEAT 1821 1832 REPEAT A.
 FT REPEAT 1833 1844 REPEAT A.
 FT REPEAT 1845 1856 REPEAT A.

FT REPEAT 1857 1867 REPEAT A (APPROXIMATE).
 FT REPEAT 1868 1879 REPEAT A.
 FT REPEAT 1880 1891 REPEAT A.
 FT REPEAT 1892 1902 REPEAT A (APPROXIMATE).
 FT REPEAT 1903 1914 REPEAT A.
 FT REPEAT 1915 1926 REPEAT A.
 FT REPEAT 1927 1938 REPEAT A.
 FT REPEAT 1939 1950 REPEAT A.
 FT DOMAIN 3536 3620 DEATH.
 FT VARSPIC 1039 1039 Q -> QFLGLHLPTAPPPLNEGESLSVRIQLQGPQTK
 FT (IN ISOFORM 2).
 FT VARSPIC 1444 3528 MISSING (IN ISOFORM 2 AND ISOFORM 3).
 FT CONFLICT 475 476 GQ -> PE (IN REF. 4).
 FT CONFLICT 971 971 I -> S (IN REF. 1).
 FT CONFLICT 3581 3582 QY -> HA (IN REF. 1).
 FT CONFLICT 3586 3586 I -> Y (IN REF. 1).
 SQ SEQUENCE 3924 AA: 430337 MW: 52AC496C428E29D2 CRC64;
 Query Match 14.9%; Score 860.5; DB 1; Length 3924;
 Best Local Similarity 29.1%; Pred. No. 3.1e-44;
 Matches 297; Conservative 147; Mismatches 404; Indels 174; Gaps 29;
 QY 3 GRKDVVEYLLQNGASVQARDGGLIPLHNACSGHAEVWVLLRHGADPNARDNWNVTPL 62
 DB 75 GHVLVQELLGRGSSVDSATKGTALHSLAGQAQAEVWVVKVKEGANINQAQSONGFTPL 134
 QY 63 HEAAIKGIDVCIVLLQHGAEPTIRNTDGR-----ALDLADPSAKAVLTGCKKDEL--- 115
 DB 135 YMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQOQHNAVALLENDTKGKVRLP 194
 QY 116 -LESARSGNEEKWALLTPLNVNCHASDGRKS-----TPLHAAAGYNRKIVQL 163
 DB 195 ALHIAARKODTKSAALLQ---NDHNADVQSKMVNRTTESGFTPLHIAAHYGNVAVTL 251
 QY 164 LLOHGADVHAKDKGLVPLHNACSYGHYEYTELLVKGACVNDMLWQFTPLHEAAASKNR 223
 DB 252 LLNRGAADVFTARGITPLHVASKRGNTNMVKLLDRLDQIDAKTRDGLTPLHCAARSGH 311
 QY 224 VEVCILLSYGADPTLLNCHNKSALDLPQPKERLAYEFKSHSLQAAREADVTRIKK 283
 DB 312 DOVVELLERGA-PLLARTKN-----GLSPLHMAAQGDHVECVK 349
 QY 284 HLSLEMVNFKHP-----QTHETALHCAASPYPKRQICELLRLKKGANINEKTEFTPL 338
 DB 350 HL-----LQHKAPVDDVDLYLTALHVA---HCGHYRVTKLLDKRANPNARALNGFTPL 402
 QY 339 HVASERAHNDVVEVVKHAKVNALDNLGOTSLHRAAYCGHLCOTCRLLSLLSYGCDPNLISL 398
 DB 403 HIACKNRKVMELLVKYASIOAITESGLTPIHVAAFMGHNLNIVLLLLQNGASPDVTNI 462
 QY 399 QGFTALQMGNNVQQLLOEGISLGNSEADROLLEAAKAGDVETVKKICTVQSVNCRDIEG 458
 DB 463 RGETALHM-----AARAGQEVVR--CLLRNGALVDARA 494
 QY 459 R-OSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNACSYGHYEYAEELLVKHGA 517
 DB 495 REEQTPHIASRLGKTEIVQLLQHMAHPDAATTNGYTPHISAREQVDVASVLEAGA 554
 QY 518 VVNVADLWKFPTPLHEAAAKGYEICKLLQHGADPTKKNRDGNTPLDLVKDGDGTDIDQLL 577
 DB 555 AHSLATKKGFTPLHVAAKYGSGLVDVAKLLQRRRAAASAGKNGLTPLHVAHIYDN----- 608
 QY 578 RGDALLDAKKGLARVKKLSPPNVNCRDTQGRHSTPLHAAAGYNNLEVAEYLLQHGA 637
 DB 609 -----QKVALLLEKGA---HATAKNGYTPHIAAKKNQMQIASTLLNYGA 653
 QY 638 DVNAQDKGLIPLHNAASYGHVDVAALLTKYNACVNATKWAFTPLHEAAQKGTOLCAL 697
 DB 654 ETNIVTKOGVTPHLASQEGHTDMVTLLLDKGANIHMTKSGLTSLHAAQEDKVNVAI 713
 QY 698 LLAHGADPTLKQEGOTPLDLV----SADVSALLTAAMPSPA-LPSCYKP----- 743
 DB 714 LTKHGADQDAHTKLGYPPLIVACHYGNVKNVNFLLKQGANVNAKTKNGYTPHQAQAQGH 773

QY 744 -QVLNGVRSPGATADALSSGSPSSLSAASSLDNLGSGFSELSSVSSSGTCASSLEK 802
 DB 774 THIIIVLLOHGAKP NATTA--NGNTAIAKRLGVI--SVVDTLKVVITEVTTTTTITE 829
 QY 803 KEVPGVDFSIQFVRLNGLEHLMDFERE---QITLDVLVEMGHKELKEIGINAYGHRHK 859
 DB 830 KHLNVPETMT-----VLVDSDDEGDDTWTGDGGEYLRPEDLKELGDDS----- 874
 QY 860 LIKGVRLISGO--OGLNPYTTLNTSGSGTILIDLSPDKKFQSVVEEMQSTVREHRDGG 917
 DB 875 -----LPSSQFLDGMN-YLRYSEGG-----RSDSLRFSFSDRSHTLSHASYLRD-- 918
 QY 918 HAGGIFRNYNLKIOKVCN-KKLERYTHRRKEVSEENHNHNERMLPHGSPFVNAIIHK 976
 DB 919 --SAVMDSDVVPISQVSTLAKAERNYRLSWGTEINDVA-----LSSSP-----IHS 966
 QY 977 GF 978
 DB 967 GF 968

RESULT 5
 ANKL_HUMAN
 ID ANKL_HUMAN STANDARD; PRT; 1880 AA.
 AC P16157;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ankyrin 1 (Erythrocyte ankyrin) (Ankyrin R).
 GN ANK1 OR ANK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.
 RC TISSUE=Hematopoietic;
 RX MEDLINE=90158830; PubMed=2137557;
 RA Lux S.E., John K.M., Bennett V.;
 RA Lambert S., Yu H., Prchal J.T., Lawler J., Ruff P., Speicher D.,
 RA Cheung M.C., Kan Y.W., Palek J.;
 RT "cDNA sequence for human erythrocyte ankyrin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).
 RN [3]
 RP VARIANT HS ILE-462.
 RX MEDLINE=96225450; PubMed=8640229;
 RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,
 RA Dornwell M., Herbers J., Kugler W., Oezcan R., Pekrun A.,
 RA Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;
 RT "Ankyrin-1 mutations are a major cause of dominant and recessive
 RT hereditary spherocytosis.";
 RL Nat. Genet. 13:214-218(1996).
 CC -!- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
 CC ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
 CC NA-K ATPASE, TO THE LYMPOCYTE MEMBRANE PROTEIN GP85, AND TO THE
 CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
 CC ERYTHROCYTE ANKIRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
 CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
 CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
 CC PLASMA MEMBRANE.
 CC -!- ALTERNATIVE PRODUCTS: At least 3 isoforms: 1/2.1 (shown here),
 CC 2/2.2 and 3; are produced by alternative splicing.
 CC -!- PTM: REGULATED BY PHOSPHORYLATION.
 CC -!- PTM: ACYLATED BY PALMITIC ACID GROUP(S).

CC -!- DISEASE: Defects in ANK1 are the cause of dominant and recessive
 CC hereditary spherocytosis (HS).
 CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.
 CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X16609; CAA34610.1; -.
 DR EMBL; M28880; AAA51732.1; -.
 DR PIR; S08275; SJHUK
 DR PIR; A35049; A35049.
 DR HSSP; Q00420; IAWC.
 DR Genew; HGNC:492; ANK1.
 DR MIM; 182900; -.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00023; ank; 24.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 22.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PSS0088; ANK_REPEAT; 20.
 DR PROSITE; PSS0297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PSS0017; DEATH_DOMAIN; 1.
 KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
 KW Phosphorylation; Lipoprotein; Disease mutation; Elliptocytosis;
 KW Polymorphism.
 FT INIT_MET 0
 FT DOMAIN 1 826
 FT FT
 FT 89 KDA DOMAIN (ANION EXCHANGE PROTEIN
 FT BINDING DOMAIN).
 FT 62 KDA DOMAIN (SPECTRIN BINDING
 FT DOMAIN).
 FT 55 KDA REGULATORY DOMAIN (REGULATES
 FT THE BINDING OF ANKYRIN TO SPECTRIN
 FT AND THE BAND 3 PROTEIN).
 FT FT
 FT REPEAT 43 72
 FT REPEAT 76 105
 FT REPEAT 109 138
 FT REPEAT 142 171
 FT REPEAT 173 200
 FT REPEAT 204 233
 FT REPEAT 237 266
 FT REPEAT 270 299
 FT REPEAT 303 332
 FT REPEAT 336 365
 FT REPEAT 369 398
 FT REPEAT 402 431
 FT REPEAT 435 464
 FT REPEAT 468 497
 FT REPEAT 501 530
 FT REPEAT 534 563
 FT REPEAT 567 596
 FT REPEAT 600 629
 FT REPEAT 633 662
 FT REPEAT 666 695
 FT REPEAT 699 728
 FT REPEAT 732 761
 FT REPEAT 765 794
 FT DOMAIN 1402 1486
 FT VARSPIC 1512 1873
 FT VARSPIC 1874 1874
 FT VARSPIC 1849 1880
 FT FT
 FT TVEGLEDPSELEVDIDYFMKSKDHTSTPNP -> ELRGS
 FT GLOPDLIGRKAQIVKRASLKRGKQ (IN ISOFORM
 FT 3).

FT DOMAIN 1387 1862
FT 55 KDA REGULATORY DOMAIN (REGULATES
FT THE BINDING OF ANKYRIN TO SPECTRIN
FT AND THE BAND 3 PROTEIN).
FT ANK 1.
FT REPEAT 40 69
FT REPEAT 73 102
FT REPEAT 106 135
FT REPEAT 139 168
FT REPEAT 170 197
FT REPEAT 201 230
FT REPEAT 234 263
FT REPEAT 267 296
FT REPEAT 300 329
FT REPEAT 333 362
FT REPEAT 366 395
FT REPEAT 399 428
FT REPEAT 432 461
FT REPEAT 465 494
FT REPEAT 498 527
FT REPEAT 531 560
FT REPEAT 564 593
FT REPEAT 597 626
FT REPEAT 630 659
FT REPEAT 663 692
FT REPEAT 696 725
FT REPEAT 729 758
FT REPEAT 762 791
FT DOMAIN 1399 1483
FT DEATH.
FT SEQUENCE 1862 AA; 204242 MW; AE6B85B5B29001E5 CRC64;
Query Match 14.5%; Score 837.5; DB 1; Length 1862;
Best Local Similarity 29.5%; Pred. No. 2.7e-43;
Matches 237; Conservative 127; Mismatches 262; Indels 177; Gaps 20;

QY 3 GRKDVVEYLQNGASQVARDGGLPLHNACSGFHGHEVNNLLRHGADPNARDNNYTP 62
DB 85 GODEVRELYNGANYNAOSKQFTPLYMAAQENHLEVVYKFLLENGANQVATEDGFTPL 144
QY 63 HEAAIKGKIDCVLLQHGAEPTI-----RNTDGRALDLADPSAKAVLTGEYK 113
DB 145 AVAQQGHENVAHLNYGFKGVRPLALHIAARNDTRTA-----AVLLQNDPNP 195
QY 114 ELLESARSGNEEKMAALLTPLNYNCHASDGKRS-----TPLHLAAGY 155
DB 196 DVL--SKTG-----FTPLHIAHYENLNVAAQLLNRGASVNFPPQNGITPLHIASRR 245
QY 156 NRKIVQLLQHGADYHAKDGLDPLHNACSYGHEVTELLVYKHCACVNMDLQFTPL 215
DB 246 GNVIMVRLLLDRGAQITRTKDELTPHCAARNGHVRISIELLDHGAQIAKTKNGLSPI 305
QY 216 HEAAKNRVVEVCSLLSYGADPTLLNCHNKSATDLAPTQPKERLAYEFKGHSLQAARE 275
DB 306 HMAAQGHDLDCVRLILQYNAE-----ID----- 328
QY 276 ADVTRIKKHLISLEWVNFKHPQTHETALHCAASPYPKRQICELLRKGANINEKTKEFL 335
DB 329 -DIT-----LDHLLTPHVA-----HCHHVRVAKVLLDKGAKPNRSLNGF 368
QY 336 TPLHVASEKAHNDVVEVYVYKAVNALDNLGOTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
DB 369 TPLHIACKKHHIRVMEILLTGTASIDAVTESGLTPHVASFMGHLPDIVKNLQRGASPNV 428
QY 396 ISLQGFALQW-----GNENYQQLQEGISLGNSEAD-----ROLLE- 432
DB 429 SNVKVETPLHMAARAGHTEVAKYLLONKAKANAKAKDDQTPHCAARIGHTGMVKLLLEN 488
QY 433 -----AAKAGDVETVKICTVQ--SVNCRDIEGROSTPLHFAAGYNRVS 474
DB 489 GASPNLATTAGTPTLHTAAREGHVDLTALALLEKEASQACMTKKG--FTPLHVAAYKGVKR 546
QY 475 VVEVLLQHGADYHAKDGLVPLHNACSYGHEVTELLVYKHCACVNMDLQFTPLH 533
DB 547 LAELLSHDHPNAAGKNGLTPLHVAHVHNNLDIVKLLPRGGSPH--SPAWNRYTPLHTA 605

QY 534 AAKGYEICKLLQHGADPTKKNRDGNTPDL-VKDGDTDIODLLRGAALLDAKKGCL 592
DB 606 AKQOIEVARSLLOYGGSANAEVQGVTPHLHAAQEGHTEWVALL----- 650
QY 593 ARVAKKLSPPDNVNCRTQGRHSTPLHLAAGYNNLEVAEYLLQHGADYNAQDKGLIPLHN 652
DB 651 -----LSKQANGNLGNKSG--LTPLHLVVSQGHVLVADVLIKHGVTVDATTRMGYTPH 703
QY 653 AASGHVDVAALLIKYNACVYNATDKWFTPLHAAQKGRTOLCALLAHAGADPTLKNOEG 712
DB 704 ASHYGNIKLVKFLQHQADVNAKTKLGYSPUHQAOQGHDTIVTLKNGASPNESVNSG 763
QY 713 QTPDL-----VSADDVSALLT 729
DB 764 TTPLAIAKRLGYISVTDVLKVV 786
RESULT 7
Y379_HUMAN STANDARD; PRT; 1059 AA.
ID Y379_HUMAN STANDARD; PRT; 1059 AA.
AC O15084;
DC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein KIAA0379 (Fragment).
GN KIAA0379.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain.
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro."
RL DNA Res. 4:141-150(1997).
CC -!- SIMILARITY: CONTAINS AT LEAST 27 ANK REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB002377; BAA20833.2; -
CC HSP; P80144; 2MYO.
CC InterPro; IPR002110; ANK.
CC Pfam; PF00023; ank; 28.
CC SMART; SM00248; ANK; 26.
CC PROSITE; PS50088; ANK_REPEAT; 24.
CC PROSITE; PS50297; ANK_REPEAT; 1.
CC KW Hypothetical protein; Repeat; ANK repeat.
FT NON_TER 1 1
FT REPEAT 46 75 ANK 1.
FT REPEAT 79 108 ANK 2.
FT REPEAT 112 141 ANK 3.
FT REPEAT 145 174 ANK 4.
FT REPEAT 178 207 ANK 5.
FT REPEAT 211 240 ANK 6.
FT REPEAT 244 273 ANK 7.
FT REPEAT 277 307 ANK 8.
FT REPEAT 311 340 ANK 9.
FT REPEAT 344 373 ANK 10.
FT REPEAT 377 406 ANK 11.
FT REPEAT 410 439 ANK 12.
FT REPEAT 443 472 ANK 13.


```
QY 44 LLRHGADPNARDNNWNYTPHLEAAIKGIDVICIVLIQHGAEPTIRNTDGRGTRALDADPSAK 103
DB 2 LLOHGAENVKADNEGRTALIAAAYNGHREIVEHLLDHGAENVHEDVDGRTALSAV-----56
QY 104 AVLTGEYKKDELLESARSNGEKKMMALITPLNVNCHASDGRKSTPLHLAAGYNRVKIVOL 163
DB 57 -----ALCPV-----ASKGHAS-----VVSL 72
QY 164 LLOHGADYHAKDKGLVPLHNACSYGHVEVTELLVKHCACVAMDLQWTFPLHEAASKNR 223
DB 73 LIDRGAENVHCHDKGWTPLVAAVEGHVDVVDLLEGGADVHTDNGRTPLLAASNGH 132
QY 224 VEVCSLLISYGADPTLLNCHNKSALDAPTPOLKERLAYEFKGHSLLO-AAREADVTRIK 282
DB 133 ASVWNTLLFWGA-----AVD-----SIDSEGRVTSIASAQNVEVVR 170
QY 283 KHLSEMNVNFKIPQTHETALHCAAAASPYPKKQICEILLRKGANTNEKTEKELPLHVAS 342
DB 171 TLLDRGLDENHRDDAGWTPHMAA---FEGRHICEALIEOGARTNEIDNDGRIPFILAS 227
QY 343 EKAHNDVVVVVYHKAQNALNDLQGTSLHRAAYCGHLQTCRLLLSYGCDPNIIISLOGFT 402
DB 228 QEGHYDCVOILLNENSNIDQRYDGRNALRVAALGHRDIVELEFSHAD-----277
QY 403 ALQMGNEVQQLLOEGISLGNSEADQRLLEAAKAGDVETVTKLCTVQSVNCRDIEGQST 462
DB 278 -----VNCKDADGRPT- 288
QY 463 PLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNACSYGHEVAELLVKHGAVVNA 522
DB 289 -LYILALENQLTMAEYFLENGANVEASDAEGRALTALVSCWQGHMEMVOVLIAHADVNA 347
QY 523 DLWKFTPLHEAAKGYETCKLLLOHGADPTKKNRDNTPDLVLKGDGTDIQLDLRGDAA 582
DB 348 DNEKRSALQSAWQGHVKKVQVLLIEHG-----A 375
QY 583 LLDAARKGLARVKKLSPDNVNCRTDGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQ 642
DB 376 VVD-----HTCNOGATALCIAAQEGHIDVQVLLHEGADPNHA 413
QY 643 DKGGLIPLHNAASYGHVDVAALLIKYNACVNTADTKWAFPTPLHEAAKQKGTQLCALLAHG 702
DB 414 DQFGRTAMRVAAKNHGSQIKLEKYGA--SSLNGCSPSPVHTMEQKPLQSLSSKV----467
QY 703 ADPTLK-NQGGOTPLDVSADVDVALLTAAMPSPALPCYKQVNLNVRSPGATADALSS 761
DB 468 QSLTIKSNSSGST-----GGGD-----MQPS-----LRGL--PNCPTHAFSS 502
QY 762 GPSPSSSL--SAASSLDNLGSGFSELSSVSSVSGTTEGASLEKKEVPGVDFSTIQFVRNL 819
DB 503 PSESPDSTVDQKSSLSNLSKSNSSRLRTTSSATAQT-----VPIDSFHNLSPTEQI 557
QY 820 GLEHLMDFEREQI 833
DB 558 QOHSPLRSRSROSI 571
RESULT 9
LATA_LATMA STANDARD; PRT; 1401 AA.
AC P23631; Q25328;
AT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-latrotoxin precursor.
OS Latrodectus mactans (Black widow spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Araneioidea; Theridiidae; Latrodectus.
OX NCBI_TaxID:6924;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tredcinguttatus; TISSUE=Venom gland;
```

```
RX MEDLINE-91031994; PubMed-1977615;
RA Kiyatkin N.I., Dulubova I.E., Chekhovskaya I.A., Grishin E.V.;
RT "Cloning and structure of cDNA encoding alpha-latrotoxin from black
RL widow spider venom."
RL FEBS Lett. 270:127-131(1990).
RN [2]
RP PARTIAL SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE-91362695; PubMed-1888339;
RA Volkova T.M., Galkina T.G., Kudelin A.B., Nazimov I.V., Grishin E.V.;
RT "Structure of tryptic fragments of a neurotoxin from black widow
RL spider venom."
RL Bioorg. Khim. 17:437-441(1991).
CC -!- FUNCTION: Binds to high-affinity receptors that are localized in
CC the presynaptic plasma membrane of the nerve terminal. Binding of
CC toxin results in calcium influx into nerve terminals, activation
CC of synaptic vesicle exocytosis, and ATP depletion.
CC -!- PTM: ALPHA-LATROTOXIN MAY BE PROCESSED AT ITS C-TERMINUS.
CC -!- SIMILARITY: CONTAINS 21 ANK REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X55009; CAA38753.1; -.
CC PTR: S11527; S11527.
CC HSP: Q00420; IAWC.
CC InterPro: IPR002110; ANK.
CC Pfam: PF00023; ank; 20.
CC SMART: SM00248; ANK; 17.
CC PROSITE: PS50088; ANK_REPEAT; 11.
CC PROSITE: PS50297; ANK_REPEAT_REGION; 1.
CC Toxin; Neurotoxin; Repeat; ANK repeat; Venom; Signal.
CC SIGNAL
CC CHAIN 1 20
CC FT REPEAT 490 521 ALPHA-LATROTOXIN.
CC FT REPEAT 525 554 ANK 1.
CC FT REPEAT 559 589 ANK 2.
CC FT REPEAT 593 622 ANK 3.
CC FT REPEAT 626 656 ANK 4.
CC FT REPEAT 660 690 ANK 5.
CC FT REPEAT 695 723 ANK 6.
CC FT REPEAT 729 758 ANK 7.
CC FT REPEAT 762 791 ANK 8.
CC FT REPEAT 795 824 ANK 9.
CC FT REPEAT 828 857 ANK 10.
CC FT REPEAT 862 891 ANK 11.
CC FT REPEAT 895 924 ANK 12.
CC FT REPEAT 928 957 ANK 13.
CC FT REPEAT 971 1003 ANK 14.
CC FT REPEAT 1004 1033 ANK 15.
CC FT REPEAT 1035 1064 ANK 16.
CC FT REPEAT 1068 1097 ANK 17.
CC FT REPEAT 1101 1131 ANK 18.
CC FT REPEAT 1137 1166 ANK 19.
CC FT REPEAT 1170 1199 ANK 20.
CC FT REPEAT 1401 AA; 156856 MW; B38A22083C142A98 CRC64;
CC SEQUENCE
Query Match 8.1%; Score 464.5; DB 1; Length 1401;
Best Local Similarity 23.2%; Pred. NO. 1.4e-20;
Matches 175; Conservative 94; Mismatches 274; Indels 211; Gaps 19;
QY 12 LONGASVQARDGGLIPLHNACSFGEAEVNVNLLRHGADPNARD-NWNTPLHEAAIKG 70
DB 513 LAKSTELNQPDKKGYTPIHVAADSGNAGIVNLLIRGVSVINSKTYHFLQTPHLAAQKGF 572
QY 71 IDVCIVLLQHGAEPTI-----RNTDGRGTRALDADPSAKAVLTGEYKKDELLESARSNGEEK 126
DB 573 VTFQRLME---SPEININERKDGFTP-----LHYAIRGERI 608
```



```

ID V222_F0MPV STANDARD; PRT; 747 AA.
AC Q9J513;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative ankryrin-repeat protein FPV222.
GN FPV222.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
CC -1- SIMILARITY: CONTAINS 14 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF198100; AAF44566.1; -
DR HSSP; Q00420; IAWC.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00023; ank; 13.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00248; ANK; 9.
DR PROSITE; PS50088; ANK_REPEAT; 7.
DR PROSITE; PS0297; ANK_REPEAT; 1.
DR Hypothetical protein; Repeat; ANK repeat.
KW REPEAT
FT REPEAT 38 67 ANK 1.
FT REPEAT 103 132 ANK 2.
FT REPEAT 136 165 ANK 3.
FT REPEAT 169 198 ANK 4.
FT REPEAT 202 231 ANK 5.
FT REPEAT 234 263 ANK 6.
FT REPEAT 294 323 ANK 7.
FT REPEAT 328 357 ANK 8.
FT REPEAT 361 393 ANK 9.
FT REPEAT 397 426 ANK 10.
FT REPEAT 430 460 ANK 11.
FT REPEAT 464 493 ANK 12.
FT REPEAT 495 524 ANK 13.
FT REPEAT 529 559 ANK 14.
SQ SEQUENCE 747 AA; 85303 MW; 55F90AF2855C3D28 CRC64;

Query Match
Best Local Similarity 6.8%; Score 394.5; DB 1; Length 747;
Matches 183; Conservative 127; Mismatches 306; Indels 193; Gaps 28;

QY 4 RKDVVEY--LLQNGASVOARDGGLPLPHNACSGEHAENVNLLRHGADPNARDNNWYTP 61
DB 16 KNDVSVRYKLEKGINPNEKNKNCNTMLXTAVEHRYDIIDIKLLDHGADPNYSSDHMT 75
QY 62 LHEAAIKGKI-DVCIVLLOHCAEPTIRNTDRTALDLADPSAKAVLTGEYKDELESAR 120
DB 76 LHSVSVIIPRKISKILTKYGNLVLIANYRNTFV-----YNEARNLEIA- 120
QY 121 SGNEEKMM-----ALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLHGHGADVHAKDK 176
DB 121 -----KMLIQNGALVNMMN-----KNITPLHIASSSGSYKMWELLLHGHGANTALTS 168
QY 177 GDLVPLPHNACSYGHEVTELLVYKHGACVNMADLWQFTPLPEHAASKNRVVCSSLLSYGAD 236
DB 169 YGETSLHYSYSSNDLNLSELLIENGTVNNVANKDSITALIIVAEINSLDLVRLLLDKGAD 228
```

```

QY 237 PTLNCHNKSALDLAPTQPKERLAYEPKGHSHLLQAAREADVTRIKKHLSLWMVNFK--- 293
DB 229 TNAIGL-ERFKLYVTETKQNNILKY-----LNTNNVNTWVTWINEYIASLELDWNRS 281
QY 294 -----HPOTHETALHCAASYPKRKQICEALLRRKAN--INEKTKEFTPLPLH 340
DB 282 ATSKLMFRSCFEPCTVPVTLATRKGS-----KELLEILLEYGCNPDIKEKTTSTVA-MHY 335
QY 341 ASEKAHNDVVEVVYKHEAKVNALNIGQTSLHRAAYCGHLQTCRLLLSYCGDPNIIISLQ 400
DB 336 AVIRKHYEMLNILIRYDAYTDVKDRQONTPAHYAVKLPISSECKYL----- 381
QY 401 FTALQMGNEVQOLLQEGISLGNSEADROLLEAAKAGDVETVKLCTVQSYNCRDIEGRQ 460
DB 382 -----KLLKLAGA-----SFNLNTRKGR- 399
QY 461 STPLHFAAGYNRVSVVEYLLQHGADVHAKDGGVLPPLHNACSYGHEVYAEVLLVKGAV-V 519
DB 400 -TPLHTACKYNTEAVKYLIESGCDTNIVDVMSFTPLNYAVYEREDTVKILLESQVDP 458
QY 520 NVADLWKFPLHEAAAKGKYEICKLLLOHGADPTKKNR-DGNTPLDLVKDGTDIQDLRL 578
DB 459 NLCDYKEVSPITQIAIKRNNKNIKMLLNAGTDIKPINECYGLHMLAALHN-----KDLL- 512
QY 579 GDAALLDAAKGGLARVKKLSSPDVNCRDQTGRHSTPL-HLAAGYNNLEVAEYLLQHGA 637
DB 513 -----KWLCTISEL-----EVNGVDD---HYPLASYVAELSDIRMEILIEKGL 555
QY 638 DVN---AQDKGGLIPLHNAASYGHVDVAALLI-----KYNACVNATDKWAF 681
DB 556 DLNKVTGPDETMTFTMFSATSDLRKSIDLLISOIAADEESEGFKINKNMIOQDKYLLR 615
QY 682 PLHEAAOKGRTQLCALLAHGADPTLNQEQGTPLDLVS-----ADVSALLTAAMPSPAL- 737
DB 616 VYHEC-----KNQVSKM-----GEIKLGDGFTMDIYKNNRSIHVNFLARYAMQLSTID 664
QY 738 -----PSCYKPOVLNGVR 750
DB 665 LREVPYIRKYLEILINPAIKRHKILNAAK 693

RESULT 12
ANR3_HUMAN STANDARD; PRT; 832 AA.
AC P57078; Q96KH0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase ANKRD3 (EC 2.7.1.-) (Ankyrin repeat
DE domain protein 3) (PKC-delta-interacting protein kinase).
GN ANKRD3 OR D1K.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Fetal kidney, and Fetal lung;
RA Shimizu N., Kudoh J., Shibuya K.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schilhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sakai T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
```

RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
 RT Lehrach H., Reinhardt R., Yaspo M.-L.;
 RL "The DNA sequence of human chromosome 21.";
 CC Nature 405:311-319(2000).
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -!- SIMILARITY: CONTAINS 10 ANK REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AB047783; BAB56136.1; -;
 DR EMBL; AP001743; BAA95526.1; -;
 DR HSSP; P25963; IIRN.
 DR Genew; HGNC:496; ANKRD3.
 DR MIM; 605706; -;
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00023; ank; 10.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR Prodom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00248; ANK; 10.
 DR SMART; SM00221; STYKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS50297; ANK_REPEAT; 9.
 DR PROSITE; PS50088; ANK_REPEAT; 9.
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;
 KW ANK repeat; Alternative splicing.
 FT DOMAIN 22 286 PROTEIN KINASE.
 FT REPEAT 485 514 ANK 1.
 FT REPEAT 518 547 ANK 2.
 FT REPEAT 551 580 ANK 3.
 FT REPEAT 584 613 ANK 4.
 FT REPEAT 617 647 ANK 5.
 FT REPEAT 651 680 ANK 6.
 FT REPEAT 684 713 ANK 7.
 FT REPEAT 717 746 ANK 8.
 FT REPEAT 750 780 ANK 9.
 FT REPEAT 782 811 ANK 10.
 FT NP_BIND 28 36 ATP (BY SIMILARITY).
 FT BINDING 51 51 ATP (BY SIMILARITY).
 FT ACT_SITE 143 143 BY SIMILARITY.
 FT VARSPIC 278 325 MISSING (IN ISOFORM 2).
 FT CONFLICT 714 714 M -> V (IN REF. 1).
 SQ SEQUENCE 832 AA; 91610 MW; 5D8FFFD5F04F7ECB CRC64;
 Query Match 6.8%; Score 390.5; DB 1; Length 832;
 Best Local Similarity 28.1%; Pred. No. 2.2e-16;
 Matches 141; Conservative 58; Mismatches 176; Indels 127; Gaps 13;
 QY 228 SLILSYGADPTLLNCHNKSADIAPTPOLKERLAYEFKGLSLQAAREADVTRIKKHLSL 287
 Db 434 SLISLSEPER-----STSDLGTTDVKKK-----LVDAIVSGDTSKLMKIL-- 474
 QY 288 EMVNFKHPQTHETALHCAAA-----SPYPRKQICELLKRGANINKEFTPLPLHVA 343
 Db 475 -----QPQVDVLALDSGASLLHLAVEAQGECAKWLILNANNPNLSNRGSGTPLHMAVE 528
 QY 344 KAHNDVVVVVVKHKAVALDNLGOTSLHRAAVCGHLQTCRLLSYGCDPNISLQGF- 402
 Db 529 RRVGVVVELLLARKISVNAKDEQWTLHFAAQNGDESSTRLLLEKNASVNEVDFFGRTP 588

QY 403 ---ALQMGNN-VOQLQEGISLGNSEADROLLEAAKAGDVETVKKLCTVQSYNCRDIEG 458
 Db 589 MHVACOHGOENIVRILLRGVDV-----SLQG 615
 QY 459 RGS-TPLHFAAGYNRVSVVEYLL-QHGADVHAKDKGGLVPLHNASYGHYEAELIVKHG 516
 Db 616 KQAWLPPLHYAAWQGHLPVVKLLAKAPGVSVNAQTLDGRTPLHLAAQGRHYRVARILIDL 675
 QY 517 AVNVADLWKFPLHEAAKGYEICKLILQHCADPTKKNRGDNTPLDLVKDGDITQDL 576
 Db 676 SDVNVCSLLAQPLHVAETGHTSTARLLHLRGAGKEAMTSDGYTALHL----- 724
 QY 577 LRGDAAALLDAAKKGLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHG 636
 Db 725 -----AARGHLATVK-----LLVEEK 741
 QY 637 ADVNAQDKGLPLHNAASYGHVDVAALIKYNACVNAATDKWFTPLHEAAQGRQTOLCA 696
 Db 742 ADVLARGPLNOTALHLAAAHGSHVEELVSADV-IDLFEQGLSALHLAAQGRHAQTVE 800
 QY 697 LLLAHGAD--PTLKNQEGQTP 715
 Db 801 TLLRHGAHINLSLQKFGGHP 822
 RESULT 13
 V244_FOWPV STANDARD; PRT; 568 AA.
 AC Q9J4Z6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Putative ankryrin-repeat protein FVP244.
 GN FVP244.
 OS Fowlpox virus (FPV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Avipoxvirus.
 OX NCBI_TaxID=10261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20193820; PubMed=10729156;
 RA Afonso C.L., Tushman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
 RT "The genome of fowlpox virus."
 RL J. Virol. 74:3815-3831(2000).
 CC -!- SIMILARITY: CONTAINS 12 ANK REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF198100; AAF44588.1; -;
 DR HSSP; P42773; I1HB.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; ank; 10.
 DR SMART; SM00248; ANK; 6.
 DR PROSITE; PS50088; ANK_REPEAT; 7.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 KW Hypothetical protein; Repeat; ANK repeat.
 FT REPEAT 40 69 ANK 1.
 FT REPEAT 144 173 ANK 2.
 FT REPEAT 177 206 ANK 3.
 FT REPEAT 210 239 ANK 4.
 FT REPEAT 272 302 ANK 5.
 FT REPEAT 306 336 ANK 6.
 FT REPEAT 340 370 ANK 7.
 FT REPEAT 374 403 ANK 8.
 FT REPEAT 407 437 ANK 9.
 FT REPEAT 441 471 ANK 10.

```
FT REPEAT 473 502 ANK 11.
FT REPEAT 571 602 ANK 12.
SQ SEQUENCE 668 AA: 76200 MW: F3C5BA076023791B CRC64;

Query Match
Best Local Similarity 27.3%; Score 380.5; DB 1; Length 668;
Matches 153; Conservative 95; Mismatches 192; Indels 121; Gaps 26;

QY 186 CSYGHVEYVELLV-----KHGA--CVNAMDLMQFTPLPHEAAAKNRVEVCSLLLSYGADPT 238
DB 11 CRNGYIELSAITNYNLKHGAQOCEN--ESIPFTAIHQALQLOIDIVKELIO--QNPX 66
QY 239 LL-----NCHNKAIDLATPOLKRLAYEFKGHSLQLOAREADVTRIKKHLSELEWFKH 294
DB 67 LIYVTDHRRNSTLHICITPNVMD-----IVISLTVDCDIILDIKYASI--ILN-KH 115
QY 295 POTHETALHCAAAASPYPRKKQICELELLLRKANINE-----KTKFELTPLHVAASEKAND-- 348
DB 116 -----KIGFACIHLKREGISGNEISYNKINKSTEYMKLIKERTQODEL 158
QY 349 -VVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLOTCRLLLSYGCDDPNIIISLQGFALQ-- 405
DB 159 LIAEMLLKKGIDVNAKDVYCRPTIHYAERGNTKMNVLNLLSYGADVNIITLDDLSVLEYA 218
QY 406 MGNENVQOLLQEGISLGN--SEADROLLEAAKAGDVETVKKLC--TVQSVNCRDIEGRQSTP 463
DB 219 VDSKNIDTITKALIDNRSNKNKNDLSLLKAIKNTDLETSLLYDSGFSVNSIDV--YKNTP 276
QY 464 LHFAAGYNRVS--VWEYLQHGADVHAKDKGLVPLHNCACSYGH--YEVAELLVKHGAVNVV 521
DB 277 LHVAVQAPSLRVLKPLKRGIDVNAKNIKGETPLYLMAKNKYDPTNITLIMRGADVNA 336
QY 522 ADLWFTPLHFAAAAGKYEICKL--LLOHGADPTKKNRGDPTPLDLVKGDDTDIDQLLRGD 580
DB 337 ADSLYITPLHQASTLDYKDTVITLLEGA----- 366
QY 581 AALLDAAKKGCLARVKLSSPDVNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADV 640
DB 367 -----NVNARDYCK--TPHYAAVNVNVVINTLLDYGDIE 402
QY 641 A-QDKGGLIPLHNAASYG---HVDVAALLIKYNACVNTDKWAFPLPHEAAKQK--RTOLC 695
DB 403 ALSQKIGTV-LH-FALYGNPYMSVKT--LIDRGANVNSKNKYLSTPLHYACKNKCKPEVI 459
QY 696 ALLLHAGADPTLKNQSGTPL 716
DB 460 KMLLDNGADVNAINIRNQPL 480

RESULT 14
V024_F0WPV STANDARD; PRT; 596 AA.
AC Q9J5H7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Putative ankryin-repeat protein FV024.
GN FV024.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP MEDLINE FROM N.A.
RX MEDLINE-20193820; PubMed-10729156;
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus."
RL J. Virol. 74:3815-3831(2000).
CC -1- SIMILARITY: CONTAINS 13 ANK REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF198100; AAF44368.1; .
CC HSSP; P42773; 11HB.
CC InterPro; IPR002110; ANK.
CC Pfam; PF00023; ank; 12.
CC SMART; SM00248; ANK; 9.
CC PROSITE; PS00088; ANK_REPEAT; 7.
CC PROSITE; PS0297; ANK_REPEAT_REGION; 1.
CC Hypothetical protein; Repeat; ANK repeat.
FT REPEAT 5 34 ANK 1.
FT REPEAT 37 66 ANK 2.
FT REPEAT 68 95 ANK 3.
FT REPEAT 99 128 ANK 4.
FT REPEAT 130 158 ANK 5.
FT REPEAT 162 191 ANK 6.
FT REPEAT 195 225 ANK 7.
FT REPEAT 229 258 ANK 8.
FT REPEAT 262 291 ANK 9.
FT REPEAT 295 324 ANK 10.
FT REPEAT 326 355 ANK 11.
FT REPEAT 359 389 ANK 12.
FT REPEAT 394 423 ANK 13.
SQ SEQUENCE 596 AA: 67938 MW: F813C43228907D07 CRC64;

Query Match
Best Local Similarity 25.2%; Score 368; DB 1; Length 596;
Matches 135; Conservative 83; Mismatches 177; Indels 140; Gaps 19;

QY 146 STPLHLAAGYNRVKIVOLLQHGADVHAKDKGLVPLHNCACSYGHVEYVELLVKHGACVN 205
DB 39 STPLYLAVNSDDIDVLFKLLDNGADI---NKCSPPHLKAINLGNVEMVKLLVDHGADIE 95
QY 206 AMDLWQFTPLHFAAASKNRVEYCSLLLSYGADPTLL-----NCHNKAIDLATPOLK 259
DB 96 KVLGN--SPLYLALCKRNTITKYLLERGADPTLFTNYCDIYKIPIDI----- 145
QY 260 LAYEFKGHSLQLOAREADVTRIKKHLSELEWFKHPOTHEHALHCAAAASPYPRKKQICE 319
DB 146 ----FK-----ILIKYKVSNIQNSHFKTPYY--AIKC---TNP-----LIK 181
QY 320 LLRKGANINERKTEFLTPHVAASEKHN---DVEVVVVKHEAKVNALDNLGQTSLHRAAY 376
DB 182 LLENNASLT--IPEGYNHHYLIITAVKHNCDISILRLIKYGVYPVNEQDLETSLSHYCVS 239
QY 377 CGHLQTCRLLLSYGCDDPNIIISLQGETALQMGNEVNVQOLLQEGISLGNSEADROLLEA 436
DB 240 AGKHDLKLLDYDADPNI-----TDSCLGTP-----LHYAVSRNDIATTLIE--KG 286
QY 437 GDV----ETVKKICTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKG 492
DB 287 ANVNIHNDIT-----DTVLTAVGNRNKILINLLLMYGANTPLKSRN 328
QY 493 GLVPLHNCACSYGHVEAELLVKHGAVNVADLWKFTPLPHEAAK--GKYEICKLLQHGAD 551
DB 329 PLI--HKAETKIDINILSEILNHGAENVINYREGVTPLYIATITFMQIKFAKLLRYG 386
QY 552 PTKKNRDNCTPLDLVKGDDTDIDQLLRGDAALLDAAKKGCLARVKKLSSPDVNVNCRDTQ 611
DB 387 PNMKN-----ES 393
QY 612 RHSTPLHLAAGYNNLEVAEYLLQHGADVNAODKGLIPLHNAASVGHVDVAALLI 666
DB 394 NENTPLHGAILSNRUDSVELLMSYNVDVHSINKLGHPTPL--SCINVISDKIATIIII 447

RESULT 15
DAPK_HUMAN
ID DAPK_HUMAN STANDARD; PRT; 1431 AA.
AC P53355;
```


QY	875	-----NPLYLNTSGGTILDLSPDDKEFFQSV EEMQSTV 910
Db	917	NRFGNDLHISNKLFVLDA GAG-----SKDKVLRNLHLOEIRS QI 956

Search completed: February 12, 2003, 03:33:01
Job time : 18.6325 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 03:26:09 ; Search time 41.3675 Seconds
(without alignments)
5478.980 Million cell updates/sec

Title: US-09-843-159B-3

Perfect score: 5769

Sequence: 1 GFGKDVVEYLLONGASVQA.....AYPEYLITYQIMRPEGMGVDG 1100

Scoring table: BLOSUM62

Capop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	3335.5	66.5	1181	5	Q9XZ37	Q9XZ37 drosophila
2	3335.5	66.5	1181	5	Q9VBP3	Q9VBP3 drosophila
3	875.5	15.2	1726	11	Q8VC68	Q8VC68 mus musculus
4	875.5	15.2	4377	4	Q12955	Q12955 homo sapien
5	873.5	15.1	1943	11	Q61307	Q61307 mus musculus
6	867.5	15.0	1136	6	Q9N180	Q9N180 bos taurus
7	864.5	15.0	2622	11	O70511	O70511 rattus norv
8	863.5	15.0	1762	11	O88521	O88521 rattus norv
9	842	14.6	1719	4	Q13768	Q13768 homo sapien
10	841.5	14.6	843	11	P97582	P97582 rattus norv
11	840.5	14.6	1856	4	Q99407	Q99407 homo sapien
12	839.5	14.6	1848	11	Q61302	Q61302 mus musculus
13	836	14.5	1549	5	Q9V4B1	Q9V4B1 drosophila
14	834	14.5	1549	5	Q24241	Q24241 drosophila
15	820	14.2	1159	5	Q9NCP8	Q9NCP8 drosophila
16	814.5	14.1	1809	5	Q17487	Q17487 caenorhabdi

17	814.5	14.1	1867	5	Q17486	Q17486 caenorhabdi
18	814.5	14.1	2039	5	Q17489	Q17489 caenorhabdi
19	810.5	14.0	1815	5	Q17488	Q17488 caenorhabdi
20	810.5	14.0	6994	5	Q17343	Q17343 caenorhabdi
21	807.5	14.0	1786	5	Q17344	Q17344 caenorhabdi
22	799	13.8	2443	5	Q9VSA2	Q9VSA2 drosophila
23	793	13.7	1009	5	O8SWY2	O8SWY2 drosophila
24	629.5	10.9	1088	4	Q13484	Q13484 homo sapien
25	628	10.9	791	4	Q9UFA4	Q9UFA4 homo sapien
26	601	10.4	1913	5	Q9GRV5	Q9GRV5 caenorhabdi
27	584.5	10.1	1619	5	O9NGR8	O9NGR8 drosophila
28	584	10.1	1453	5	Q9VMR4	Q9VMR4 drosophila
29	554.5	9.6	1098	11	Q61304	Q61304 mus musculus
30	545.5	9.5	1411	5	Q02989	Q02989 latrodectus
31	534.5	9.3	1599	11	Q99NH0	Q99NH0 mus musculus
32	527	9.1	743	4	Q8TB46	Q8TB46 mus musculus
33	520	9.0	1395	5	O9XZC0	O9XZC0 latrodectus
34	519	9.0	1282	5	O8T4F8	O8T4F8 drosophila
35	519	9.0	2119	5	O9VAU5	O9VAU5 drosophila
36	519	9.0	2119	5	O8T9G9	O8T9G9 drosophila
37	514	8.9	934	16	O83807	O83807 treponema p
38	509	8.8	1021	13	Q8UVC1	Q8UVC1 brachydanio
39	500.5	8.7	1486	4	Q8TEF1	Q8TEF1 homo sapien
40	499.5	8.7	833	4	Q96I86	Q96I86 homo sapien
41	496.5	8.6	1062	11	O89019	O89019 mus musculus
42	491.5	8.5	1062	11	O88849	O88849 mus musculus
43	489.5	8.5	1188	4	Q9H288	Q9H288 homo sapien
44	486	8.4	895	4	Q9Y488	Q9Y488 homo sapien
45	486	8.4	1065	4	Q9Y283	Q9Y283 homo sapien

ALIGNMENTS

RESULT 1

Q9XZ37 ID Q9XZ37 PRELIMINARY: PRT: 1181 AA.
AC Q9XZ37;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Hypothetical 127.9 kDa protein.
GN BCDNA:LD22548 OR CG4719 OR CGI7487.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Abayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenhoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Gallego R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svirska R.R., Weinburg T.,
RA Celnikier S.E.
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF132196; AAD34784.1; -.
DR HSSP: Q00420; IAWC.
DR FlyBase: FBGN0027508; BCDNA:LD22548.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR001660; SAM.
DR Pfam: PF00023; ank; 17.
DR Pfam: PF00536; SAM; 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 15.
DR SMART: SM00454; SAM; 1.
DR PROSITE: PS50088; ANK_REPEAT; 14.
DR PROSITE: PS50297; ANK_REPEAT; 1.
KW ANK repeat; Hypothetical protein; Repeat.
SQ SEQUENCE 1181 AA; 127884 MW; 4E2BF31549BBA2D9 CRC64;

Query Match	66.5%	Score	3835.5	DB	5	Length	1181
Best Local Similarity	65.3%	Mismatches	7.4e-237				
Matches	740	Conservative	156	Gaps	182	Indels	55
1	GFGRKDVVEYL	L	L	L	L	L	L
66	GYGRREVEFL	L	L	L	L	L	L
61	PLHEAAIKGKIDV	L	L	L	L	L	L
126	PLHEAASGKVDV	L	L	L	L	L	L
121	SGNEEKWALLT	L	L	L	L	L	L
186	SGAEDRLIALT	L	L	L	L	L	L
181	PLHNACSYGHV	L	L	L	L	L	L
246	PLHNACSYGHF	L	L	L	L	L	L
241	NCHNKSATDAP	L	L	L	L	L	L
306	NCHSKSAIDAA	L	L	L	L	L	L
301	ALHCAAASPYK	L	L	L	L	L	L
366	PLHLAVVSPD	L	L	L	L	L	L
361	NALDNLGOTS	L	L	L	L	L	L
426	NALDSLGOTPL	L	L	L	L	L	L
421	LGNSEADROL	L	L	L	L	L	L
480	-NPPDSETHL	L	L	L	L	L	L
479	LLOHGADYHAK	L	L	L	L	L	L
539	LLEHGAEVYAA	L	L	L	L	L	L
539	YEICKLLOHG	L	L	L	L	L	L
599	YDICKLLUHG	L	L	L	L	L	L
599	SSPDNVNCRD	L	L	L	L	L	L
659	VTPSEINCRDA	L	L	L	L	L	L
659	VDVAALLIKY	L	L	L	L	L	L
719	LDIAALLIKHT	L	L	L	L	L	L
719	VSADDVSALL	L	L	L	L	L	L
779	ATADDVKLLO	L	L	L	L	L	L
773	SSLDNLSGFS	L	L	L	L	L	L
838	TETVLLPTCA	L	L	L	L	L	L
814	OFVRNLGLH	L	L	L	L	L	L
896	GFLSSQQULH	L	L	L	L	L	L
874	LNPLYTLMT	L	L	L	L	L	L
952	-----TTGI	L	L	L	L	L	L
928	ILKITKVCN	L	L	L	L	L	L
1004	IIRQVQNRKL	L	L	L	L	L	L
988	MFAGGIYE	L	L	L	L	L	L

[illegible]

Qy	928	ILKTKVCNKKLWERYTHRRKKEVSENNHNHNERMLFHGSPFVNALIHKGFDERHAYTGG	987
Db	1004	IIRVQVQNRLKWERVAHRQRQETAEENFLQSNRMLFHGSPFVNALIHRGFDERHAYTGG	1063
Qy	988	MFGAGIYFAENSSKSNQYVYIGGGTGCPCVHKDRSCYICHRQLLCRVTTLGKSKFLQFSAM	1047
Db	1064	MFGAGIYFAHSSKSNQYVYIGGGTGCPCVHKDRSCYICHRQLLCRVTTLGKSKFLQFSAM	1123
Qy	1048	KMAHSPGHHHSVTGRPSVNGALAAEYVIYRGEQAYPEYLITQIMRPEGWGD	1100
Db	1124	KMAHAPGHHHSVGRPSAGGLHFAEYVYRGEQYPEYLITQIVKPDSSSG	1176
<p>RESULT 3</p> <p>Q8VC68 ID Q8VC68 PRELIMINARY; PRT: 1726 AA.</p> <p>AC Q8VC68</p> <p>AD Q8VC68</p> <p>DT 01-MAR-2002 (TrEMBLrel. 20, Created)</p> <p>DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)</p> <p>DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)</p> <p>DE Hypothetical 188.2 kDa protein.</p> <p>OS Mus musculus (Mouse).</p> <p>OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;</p> <p>OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.</p> <p>OX NCBI_TaxID=10090;</p> <p>RN [1]</p> <p>RP SEQUENCE FROM N.A.</p> <p>RC TISSUE=Salivary gland;</p> <p>RA Strausberg R.;</p> <p>RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.</p> <p>DR EMBL; BC021657; AAH21657.1; -</p> <p>DR InterPro; IPR002110; ANK.</p> <p>DR InterPro; IPR000488; Death.</p> <p>DR InterPro; IPR000906; ZU5.</p> <p>DR Pfam; PF00023; ank; 24.</p> <p>DR Pfam; PF00531; death; 1.</p> <p>DR Pfam; PF00791; ZU5; 1.</p> <p>DR PRINTS; PR01415; ANKYRIN.</p> <p>DR SMART; SM00248; ANK; 22.</p> <p>DR SMART; SM00005; DEATH; 1.</p> <p>DR SMART; SM00218; ZU5; 1.</p> <p>DR PROSITE; PS50088; ANK_REPEAT; 21.</p> <p>DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.</p> <p>DR PROSITE; PS50017; DEATH_DOMAIN; 1.</p> <p>DR Hypothetical protein.</p> <p>SK SEQUENCE 1726 AA; 188239 MW; 430148B67E00B1CA CRC64;</p>			
<p>Query Match 15.2%; Score 875.5; Db 11; Length 1726;</p> <p>Best Local Similarity 28.7%; Pred. No. 5e-47;</p> <p>Matches 320; Conservative 158; Mismatches 465; Indels 171; Gaps</p>			
Qy	3	GRKDVEYLLONGASVOARDGGGLIPLHNACSGHAEVYNLLRHGADPNARDNNYTP	62
Db	68	GHVEVSELLQREANVDATKKGNTALHIASLAGQAEVVKLVNTGANYNAQSQNGFTPL	127
Qy	63	HEAAIKGKIDVICVILLOHGAEPRTWDTGDTALDLA-----DPSKAVLTGEYK---KDE	114
Db	128	YMAQENHLEVYRFLDNGASQSALATEDGFTPLAVALAQOHQDVVSLLENDTKGKVRLP	187
Qy	115	LLESARSGNEKEMALLTPTLVNCHASDGRKSTPLHLAGYNRVKTVOLLQHGADVHAK	174
Db	188	ALHTAARKDQTKAAALLLQNDTNADVESKSGFTPLHIAHYGNIIVATLLNRAAVDFT	247
Qy	175	DGDLVPLHNACSYGVHYETVELLVKKGACVNMADLWQFTPLHEAAKSNRVEVCSLLLSY	234
Db	248	ARNDITPLRVASKRGANMYKLLLDGAKIDAKTRDGLTPLHCGARSGHEQVVEMLD	307
Qy	235	ADPTLLNCHNKSAIDLAPTQPKERLAYBEKGHSLLOQAREADVTRIKHLSLEMVNFKH	294
Db	308	A-PILSKTKN-----GLSPHLMATQGD-----HLNCVOLLLOH	339
Qy	295	-----POTHEIALHCAASPYKPKRQICEILLRKGANIKEKTFTPLHVASEKAHN	347

Db 340 NVPDVTNDYLTALHVAHCCHYK---VAKVLLDKKASPNKALNGFTPLHIACKKNRI 396
QY 348 DVVEVVVHKEAKVNDLNGOTSLHRAAYCGHLCOTRLLSYGCDPNITISLOGFTALQMG 407
Db 397 RVWELLKHGASIQAVTESGLTPIHVAAPMGHVNIVSQLMHGASPNITNVRGETALHMA 456
QY 408 -----NENVOQLLOEGISLGNSEADRO---LLEAAKAGDVETVKKLCCTV-QSVNCRDIEGR 459
Db 457 ARSGQAEVRYLVQDGAQVEAKAKDDQTPHLSARLGRADIVQQLLOQASPNATTS- 515
QY 460 QSTPLHFAAGYNRVSVVEYLLOHGADVHAKDGLVPLHNAACSYGHEVEAEVLLKHGAVV 519
Db 516 -YTPHLAAREGHEDEVAFLLDHGASLSITTKKGTPLHVAAYKGLVAVALLQKSASP 574
QY 520 NVADLWKFPLHEAAKGYEICKLLLOHGADPTKKNRDNTPDL-LVKDGTDDIQLL- 577
Db 575 DAAGSGSLTPLHVAHYDNQVALLLLOQASPHAAKNNGYTPHIAAKKNOMDITATSL 634
QY 578 -----RGDAALLDAKKGCLARYKLLSSPD-NVNCRDGTGRHSTPLHLAAGYNN 625
Db 635 EYGADANAVTROGIAVSHLAQEGHVDVMVSLLSRNANVLSNKS- -LTPHLAAQEDR 692
QY 626 LEVAYEYLLOHGADVNAQKGGILPLHNAASYGHVDVAALLIKYNACVNAATDKWATPLHE 685
Db 693 VNAEVLVNOGAHVAQTKMGYTPPHVGHYGNIKIVNFFLQHSKAVNAKTKNGYALTQ 752
QY 686 AAKGRTOLCALLAHGADPTLKNQEGOTPLDLVS-----ADVSALLTAAMPSPALP 738
Db 753 AAQOQHTHIINVLONNAPSNELTVNGVNTALAIARRLGYSIVVDITKVVTEIMTTIT 812
QY 739 SCYK---POVLNGV--RSPGATADLSS-----GP-----SSPSLSAASSLDNL 778
Db 813 EKHKNVPTMNEVLMDSDDEDAITGTDKYLGPQDLKELGDDSLPAEGYVGSGLGAR 872
QY 779 SCSFSELSSVVS-----SSGTGASLEKKEVPGVDFSTIQVRLNGLHLMDFIERQI 833
Db 873 SASLSFSDSRSTYLRNSYARDSMMIBELLVPSKEQHLT-FTREFDSDSLR-YSWAAD 930
QY 834 TLVDL-----VEMGHKELKEIGINAYG-----HRHKLIT-----KGVRLIS 869
Db 931 TLDNVLNVSPPVHSGF--LVSPMVDARGSMGRHGHMIRIIPRCKTAPTRITCLRVK 988
QY 870 GOOGLNPLYLTNTSGSTLIDLSPDKEFQS-VEEEMQSTVREHRDGGHAGGIFNRNYI 928
Db 989 RHKLANPPPMVEGELASRLAVEMGPAGAQLGPIVIEIP-----HFGSMRGKERE 1038
QY 929 LKIOKVCNKKLWE--RYTHRRKEVE-----ENHNHANRMLFHGSPFVNAIHKGEDER 981
Db 1039 LIVRSEGETWKEQFDSKNEIDLAEELNGMDELDSPLEEL---GTRICRIITKDFPQ- 1094
QY 982 HAYIGMFGAGIYFAENS---SKSQYVYVIGGG 1012
Db 1095 -----YFAVVSRIKQESNQ--IGPBG 1114

RESULT 4
Q12955
ID Q12955 PRELIMINARY; PRT: 4377 AA.
AC Q12955
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Ankyrin G.
GN ANK-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN STEM;
RX MEDLINE=95138209; PubMed=7836469;

RA Kordeli E., Lambert S., Bennett V.;
RT "AnkyrinG. A new ankyrin gene with neural-specific isoforms localized
RL at the axonal initial segment and node of Ranvier.";
RN J. Biol. Chem. 270:2352-2359(1995).
RP [2]
RC SEQUENCE FROM N.A.
RA TISSUE=BRAIN STEM;
RC Carpenter S.S.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U13616; AA64834.1; -;
DR HSP; P55273; IBI8.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS00088; ANK_REPEAT; 21.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 4377 AA; 480395 MW; F42379E55768B684 CRC64;
Query Match 15.2%; Score 875.5; DB 4; Length 4377;
Best Local Similarity 30.5%; Pred. No. 2.1e-46;
Matches 273; Conservative 126; Mismatches 362; Indels 135; Gaps 21;
QY 3 GRKDVYVLLQNGASVQARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNWYTP 62
Db 85 GHVEVVELLQREANVDAATKKGNTALHIALAGAEVVKVLTNGANVNAQSQNGFTPL 144
QY 63 HEAAIKGIDVCIIVLLOHGAETPIRTDGRYALDLA-----DPSKAVUTGEYK---KDE 114
Db 145 YMAAGENHLEVVKFLLDNGASQSLATEDEGFTPLAVALQOQHDQVVSLLLENDTKGVRLP 204
QY 115 LLESARSGNEEKMMALPLNVNCHASDGRKSTPLHLAAGYNRKIVOLLLOHGADVHAK 174
Db 205 ALHIAARKDDTKAAALLLQNDNADVESKSGFTPLHIAHYGINIVATLLNRAAVDT 264
QY 175 DKGDVLPLHNACSYGHEVEYVELLVKHGACVNAQMLQVTPPLHEAAKSNRVEVCSLLSYC 234
Db 265 ARNDITPLHVASKRGNNANMVKLLDRGAKIDAKTRDGLTPLHCGARSGHEQVVMELDLRA 324
QY 235 ADPTLLNCHNKSAIDLATPQLKERLAYEFKSHLLOAREADVTRIKKHLSELMVNFH 294
Db 325 A-PILSKTKN-----GLSLPHMATQGD-----HLNCVQLLLOH 356
QY 295 -----PQTHETALHCAASPYPKRKOICEILLRKGANINEKTEFTPLHVAEKAHN 347
Db 357 NVPDVTNDYLTALHVAHCCHYK---VAKVLLDKKASPNKALNGFTPLHIACKKNRI 413
QY 348 DVVEVVVHKEAKVNDLNGOTSLHRAAYCGHLCOTRLLSYGCDPNITISLOGFTALQMG 407
Db 414 KVMELLKHGASIQAVTESGLTPIHVAAPMGHVNIVSQLMHGASPNITNVRGETALHMA 473
QY 408 -----NENVOQLLOEGISLGNSEADRO---LLEAAKAGDVETVKKLCCTV-QSVNCRDIEGR 459
Db 474 ARSGQAEVRYLVQDGAQVEAKAKDDQTPHLSARLGRADIVQQLLOQASPNATTS- 532
QY 460 QSTPLHFAAGYNRVSVVEYLLOHGADVHAKDGLVPLHNAACSYGHEVEAEVLLKHGAVV 519
Db 533 -YTPHLAAREGHEDEVAFLLDHGASLSITTKKGTPLHVAAYKGLVAVALLQKSASP 591
QY 520 NVADLWKFPLHEAAKGYEICKLLLOHGADPTKKNRDNTPDL-LVKDGTDDIQLL- 577
Db 592 DAAGSGSLTPLHVAHYDNQVALLLLOQASPHAAKNNGYTPHIAAKKNOMDITATSL 651
QY 578 -----RGDAALLDAKKGCLARYKLLSSPD-NVNCRDGTGRHSTPLHLAAGYNN 625


```
QY 115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAK 174
DR HSSP: Q00420; 1AWC.
DB 188 ALHIAARKDDTRAAALLLQNDNADVESKSGFTPLHIAHYGNINVATLLNRAAVDT 247
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000906; ZU5.
DR Pfam: PF00023; ank; 24.
QY 175 DKGDLVPLINACSYGHEVTELLVHGACVNDLWQFTPLHEAASKNRVVCSSLLSYG 234
DR Pfam: PF00791; ZU5; 1.
DB 248 ARNDITPLHVASKRGNNANMVKLLDGRGAKIDAKTRDGLTPLHCGARSQEVVEMLLDS 307
DR SMART; SM00248; ANK; 22.
DR SMART; SM00218; ZU5; 1.
QY 235 ADPTLLNCHKNKAIDLAPTPQLKERLAYEFKGSHLSLLOAREADVTRIKHLSLEWVNFH 294
DR PROSITE; PS50088; ANK_REPEAT; 20.
DB 308 A-PILSKTKN-----GLSPHMTAQD-----HLNCVQLLQH 339
DR PROSITE; PS50297; ANK_REPEAT; 1.
QY 295 -----POTHETALHCAASPPKRCQICELLRLKRGANINEKTEFLPLHVASEKAHN 347
DR ANK repeat; Repeat.
DB 340 NVPVDVDTNDYLTALHVAHCHGYK---VAKVLLDKASPNKALNGFTPLHIACKNRI 396
DR ANK repeat; Repeat.
QY 348 DVVEVVVHKAQVNDLNGQTSRLHRAAYCGHLCRLLSYGCDPNIISLQGTALOMG 407
DR ANK repeat; Repeat.
DB 397 RVWELLKHGASIQAVTESGLTPIHVAAPFMGHVNIIVQLMHGASPTTNRGETALHMA 456
DR ANK repeat; Repeat.
QY 408 -----NENVOOLLOBGISLGNSEADQ---LLEAAKAGDVETVKKLCIV-QSVNCRDIEGR 459
DR ANK repeat; Repeat.
DB 457 ARSGQAEVRYLVQDGAQVEAKAKDDQTPHLSARLGKADIVQQLLQOGASPNAAITSG- 515
DR ANK repeat; Repeat.
QY 460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKGAVV 519
DR ANK repeat; Repeat.
DB 516 -YTPLHLAAREGHEDVAAFLLDHGASLSITTKGFTPLHVAAYKGLVASLLLOKSASP 574
DR ANK repeat; Repeat.
QY 520 NVADLWKFPLHEAAKAGYEICKLLOHGADPTKKNRCDNTPLDL-VKGDGTDIDOLL- 577
DR ANK repeat; Repeat.
DB 575 DAAGKSGLTPLHVAAYHYNQKVALLLDQASPHAAKNGYTPHIAAKKNQMDIATSL 634
DR ANK repeat; Repeat.
QY 578 -----RCDALDLAAKGCCLARYKKLSSPD-VNCRDTCGRHSTPLHLAAGYNN 625
DR ANK repeat; Repeat.
DB 635 EYGADANAVTROGIAVHUAQEGHVDVMSLLSRNANVNLNKGSG-LTPLHLAAQEDR 692
DR ANK repeat; Repeat.
QY 626 LEVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIRYNACVNATDKWAFPLHE 685
DR ANK repeat; Repeat.
DB 693 VNVAEVLVNOGAHVDAQTKMGVYTPHVGCHYGNIKIVNELLQHSKAVNAKTNGYTLHQ 752
DR ANK repeat; Repeat.
QY 686 AAKGRTQLCALLAHGADPTLKNQEGQTPDLVS-----ADVSALLTAMPSPALP 738
DR ANK repeat; Repeat.
DB 753 AAQOGHTHIINLQNNASPNELTVNGNTALAIARLGYISVVDLTKVYTEIMTTIT 812
DR ANK repeat; Repeat.
QY 739 SCYK---POVLNGVRSPGATADALSGSPSSLSAASLDNLS-GSFSELSSVVSSTGT 794
DR ANK repeat; Repeat.
DB 813 EKHKNVPTMNEVLD-----MSDEVKASAPEKLSGEY-----ISDGE 853
DR ANK repeat; Repeat.
QY 795 EGASSLEKKEVPCVDSITQFVRNLGLEHMDIFEREQITLDVLVEMGHKELKEIG 850
DR ANK repeat; Repeat.
DB 854 EG-----EDAITGDTKYLGPQDLKELG 876
DR ANK repeat; Repeat.
RESULT 6
Q9N180
ID Q9N180 PRELIMINARY; PRT; 1136 AA.
AC Q9N180;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Ankyrin 1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRED JAPANESE BLACK; TISSUE=BONE MARROW;
RA Matsumoto M., Inaba M., Koshino I., Saito D., Ono K.;
RT "Cloning of bovine erythrocyte ankyrin.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; AF222766; AAF61702.1; -.
DR HSSP: Q00420; 1AWC.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000906; ZU5.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00791; ZU5; 1.
DR SMART; SM00248; ANK; 22.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat; Repeat.
FT NON_TER 1136 AA; 123182 MW; 07CD8B1568ACE8C7 CRC64;
SQ SEQUENCE 1136 AA; 123182 MW; 07CD8B1568ACE8C7 CRC64;

Query Match 15.0%; Score 867.5; DB 6; Length 1136;
Best Local Similarity 26.6%; Pred. No. 8.7e-47;
Matches 283; Conservative 157; Mismatches 375; Indels 249; Gaps 28;

QY 3 GRKDVVEYLLQNGASVQARDGGLIPLHNACSYGHEVVAELLVKGACVNDLWQFTPL 62
DB 88 GODEVVRVRELNVYGANVNAQSQGFTPLYMAQENHLEVVKFLLENGANQNVATEDGFTPL 147
QY 63 HEAAITKGITDVCIVLLOHGAETPI-----RNTDGTALDLADPSAKAVUTGEYKDK 113
DB 148 AVALOQGHENVVAHLINYGTKGVRPLPALHIAARNDDTRTA-----AVLLQNDPNP 198
QY 114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS-----TPHLHLAAGY 155
DB 199 DVL--SKTG-----FTPLHIAHYENLNVAAQLLNNGASVNFPPQNGITPLHIASRR 248
QY 156 NRVKIVQLLQHGADVHAKDKGDLVPLHNACSYGHEVTELLVKGACVNDLWQFTPL 215
DB 249 GNVIVVRLLLDRGAQIETRTKDELTPHCAARNHGLRISEILLDHGCAQIAKTNGLSPI 308
QY 216 HFAASKNRVVECSLLSYG---DPTL-----LNC-HNKA---IDLAPTOLKER 259
DB 309 HMAAQGDHLDCVRLLLQYNAEIDITDLTPLHVAAGHGHRRVAKVLLDKGAKPNRS-- 366
QY 260 LAVEFKGSHLSLLOAAREADVTRI-----KKHLSLEWVNFKHPOTHETALHCAASVYPKR 314
DB 367 ---ALNGFTPLHIACKNHIRVWELLKMGASIDAVT-----ESGLTPLHVAFSPGHP--- 416
QY 315 QICELLRLKGANINEKTEFLTPLHVAASEKANDVVEVVVVKHEAKYNALDNLGQTSLHRA 374
DB 417 PIVKSLLOREASPNVSNVQVETPLHMAARAGHTEVAKYLLQNKAKYNAKAKDDQTPHLCA 476
QY 375 AYCGLHQTCLRLLSYGCDDNIISLQGTALOMGNENVOQLLQEGISLGNSEADRLLEAA 434
DB 477 ARIGHTNMVYKILLNANPNLATAGTTPLHI-----AA 510
QY 435 KAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKG 493
DB 511 REGHVETALALLEKESQTCMTKKG--FTPLHVAAYKGVKRMALLEHDAHHPNAAGKSG 568
QY 494 LVPLHNACSYGHEVVAELLVKGAVVNVADLWQFTPLHEAAAKGKYEICKLILQHGADPT 553
DB 569 LTPLHVAVHHNHLVDVRLLLPRGSSPHSPALNGYTPHIAAQNLQLEVAARSLLYQGSAN 628
QY 554 KKNRGNTPLDL-VKGDGTDIDQLLRGDALLDAKKGCLARVKKLSSPDNVNCRDTCGR 612
DB 629 AESVQGVTPHLHAAQEGHAEWVALL-----LSKQANGNLGNKSG- 667
QY 613 HETPLHLAAGYNNLEVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACV 672
DB 668 -JTPLHVAQEGHIPVADVLIKGVTVDTATRMGYTPHVAHYGNIKLVKFLQLHKADV 726
QY 673 NATDKWAFPLHEAAKAGRTQLCALLAHGADPTLKNQEGQTPDL-----VSADDVSA 726
DB 727 NAKTKLGYSPLHQAQOQGHDTIVTLLKKGASPNVSSNGTTPLAIAKRLGYISVTDVLK 786
QY 727 LITAAMPSPALPSCYK---POVLN-----GYRSPGATA----- 756
DB 787 VVTDETTVOLITDKHRMSPFETVDEITLDSDEGEDGLGPRAPERDARDVEEKELDFVP 846
```

QY 757 --DALSSGSPSP-----SSLSAASLDNLGSGFSSELSSVSSSGTEGASSLEKKEVP 806
 Db 847 KLDQVSPSPALPRICVPTETVWIRADPDQVSKYEYDESLIPSPATETSDNISPVASP 906
 QY 807 GVDFSITOFVRLNGLEHLMDFEREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVE- 865
 Db 907 VHTGFLVSPF-----VDARGGSMRGSRHRLGRVVVP 937
 QY 866 -----RLISGOGLNPYLTLNTSGSTLIDLSP----- 894
 Db 938 PRTCAAPTRICRULVKPKLPTPPPLAEELGASVIALGPTGVQFLSPVIVPEIPHFASQ 997
 QY 895 --DDKEFOSVEEMOSTVREHRRDGGHAGGIFNRYNLIKIQKVCN 936
 Db 998 GRGDRELVVLRSENGSVNKEH-----NRYGESYLDQILN 1032
 RESULT 7
 O70511
 ID O70511 PRELIMINARY; PRT: 2622 AA.
 AC O70511;
 DT 01-AUG-1998 (TremBLrel. 07, Created)
 DT 01-AUG-1999 (TremBLrel. 11, Last sequence update)
 DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
 DE 270 kDa ANKYRIN G isoform (ANKYRING) (Fragment).
 GN ANK3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP MEDLINE=98417633; PubMed=9744885;
 RA Zhang X., Bennett V.;
 RT "Restriction of 480/270-kD ankyrin G to axon proximal segments
 requires multiple ankyrin G-specific domains.";
 RL J. Cell Biol. 142:1571-1581(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Carpenter S.S., Zhang X.;
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 934-1220 FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SKELETAL MUSCLE;
 RX MEDLINE=98330444; PubMed=9664041;
 RA Kordeli E., Ludosky M.A., Deprette C., Frappier T., Cartaud J.;
 RT "Ankyrin G is associated with the postsynaptic membrane and the
 sarcoplasmic reticulum in the skeletal muscle fibre.";
 RL J. Cell Sci. 111:2197-2207(1998).
 DR EMBL; AF102552; AAC78143.1; -;
 DR EMBL; AF065150; AAC18853.1; -;
 DR HSP; P55273; 1B18.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR004888; Death.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00023; ank; 24.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00791; ZU5; 1.
 DR SMART; SM00248; ANK; 21.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 21.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 KW ANK repeat; Repeat.
 FT NON_TER 2622 2622
 SQ SEQUENCE 2622 AA; 284456 MW; 67B34830D3AC884E CRC64;

Query Match 15.0%; Score 864.5; DB 11; Length 2622;
 Best Local Similarity 30.5%; Pred. No. 4.8e-46;
 Matches 276; Conservative 123; Mismatches 360; Indels 147; Gaps 23;

QY 3 GRKDVVEYLQNGASVQARDGGGLIPLHNACSGFHAEEVYNLLRHGADPNARDNNMYTPL 62
 Db 85 GHVEVVELLOREANVDAATKKGNTALHTASLAGOAEVVKVLTNCANVAOQSNGFTPL 144
 QY 63 HPAALTKGTDVCIIVLQGAETIRNTDGTALDLA-----DPSKAVLTGEYK---KDE 114
 Db 145 YMAAQENHLEVVYRFLIDNGASQSLATEDGFTPLAVALOQGHQVLSLLLENDTKGVRLP 204
 QY 115 LLESARSGNEEKMMALL-----TPLNVNCHASDGRKSTPLHLAAGYNNRVKIVQL 164
 Db 205 ALHIAARKDDTKAALLLQNDTNADIESKMNVRNATESG--FTSLHIAAHYNINATLL 262
 QY 165 LOGADVHAHAKGDLVPLHNACSYGHEYTELLVHKGACVNMADLWQFTPLHEAASKNRV 224
 Db 263 LNKAAAVDFETNDITPLHVASKRGNNAMVKLLDRGAKIDAKTRDGLTPLHCGARSHE 322
 QY 225 EYCSLLLSYGADPTLLNCHINKSAIDLAPTPOLKERLAYEEKHSLLOAAAREADVTIKKH 284
 Db 323 QVEMLLDRAA--PILSKTKN-----GLSPUHMATQGD-----H 354
 QY 285 LSELMVNFKH-----POTHETALHCAASPYPRKRKQICEILLRKGANINEKTKFLTP 337
 Db 355 LNCVOLLLOHNVVDVTDNYLTALHVAHCHGYK---VAKVLLDKKANPNKALNGFTP 411
 QY 338 LHVASEKAHNDVVEVVKHEAKVNALDNLGQTSLHRAAYCGHLOTCLRLLSYGCDDNIIS 397
 Db 412 LHIACKNRIRVWELLKKGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHGASPNNTN 471
 QY 398 LOGFTALQMG-----NENYQQLLOEGISLGNSEADRQ--LLEAAKAGDVETVKKLTCTV-Q 449
 Db 472 VRGETALHMAARSQGAEEVRYLVODGAQVEAKAKDDQTPLHISARLKGADIVQOLLQOGA 531
 QY 450 SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGLVPLHNACSYGHEVA 509
 Db 532 SPNAATTSG--YTPHLHSAREGHEDVAAFLLDHGCASLSITTKKGFTPLHVAAYGKLEVA 589
 QY 510 ELLVHKGAVNVVADLWKFTPLHEAAGKYEICKLLQHGADPTKKNRGNTPLDL-VKD 568
 Db 590 SULLQKASPDAAAGKSGLTPLHVAHYDNQKVALLLLDQGCASPHAAAKNGYTPLHIAAKK 649
 QY 569 GUTDIQDLL-----RGDAALLDAKKGCLARVKKLSSPD--NVNCRDTQGRHST 615
 Db 650 NQMDJATSLLEYGADANPVYTRGIVASHVLAQEGHVDVMSLLSRNANVNLNKSQ--LT 707
 QY 616 PLHAAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALIKYNACVNA 675
 Db 708 PLHAAQEDRVNVAEVLVNOGAHVDAQTKMGYTPLVHVGCHYGNLIKIVNPLQHSKVN 767
 QY 676 DKWAFPLHEAAQKGRQTQCALLLAHGADPTLKNQGGQTPDLVS-----ADVSALL 728
 Db 768 TKNGYTPLHQAQOQGHTHIINVLLQNNASPNELTVNGNTALAIARRLGYISVVDTLKVV 827
 QY 729 TAAMPSPALPSYK--POVLNGVRSRPGATADALSSGSPSSLSAASLDNLS-GSFSE 784
 Db 828 EEIMTTTITERKHKNVPEPMNEVD-----MSDDEVGKASAEKLSDEY-- 873
 QY 785 LSSVSSSGTEGASSLEKKEVPGVDFSIQFVRNLGLEHLMDFEREQITLDVLVEMGHK 844
 Db 874 -----ISDGEQ-----EDAITGDTDKYLGQP 895
 QY 845 ELKEIG 850
 Db 896 DLKELG 901
 RESULT 8
 O88521
 ID O88521 PRELIMINARY; PRT: 1762 AA.
 AC O88521;
 DT 01-NOV-1998 (TremBLrel. 08, Created)
 DT 01-NOV-1998 (TremBLrel. 08, Last sequence update)
 DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
 DE 190 kDa ankyrin isoform.

```

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98395113; PubMed=9727010;
RA Thevananthar S., Kollu A.H., Devarajan P.;
RT "Identification of a novel ankyrin isoform (AnkG190) in kidney and
RT lung that associates with the plasma membrane and binds alpha-Na, k-
RT ATPase.";
RL J. Biol. Chem. 273:23952-23958(1998).
DR EMBL; AF069525; AAC34809.1; -.
DR HSP; P55273; IBI8.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000169; Shprot_acsite.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZUS; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 19.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; UNKNOWN_1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1762 AA; 191899 MW; 9023280086A7BF4E CRC64;

Query Match          15.0%; Score 863.5; DB 11; Length 1762;
Best Local Similarity 30.2%; Pred. No. 3.1e-46;
Matches 270; Conservative 123; Mismatches 369; Indels 133; Gaps 21;

Qy 3 GRKDVVEYLLQNGASVQARDGGLPLHNACSFHAEVYVNLRLRHGADPNARDNNYTP 62
Db 79 GIVEVSEYLLQNGASVQARDGGLPLHNACSFHAEVYVNLRLRHGADPNARDNNYTP 138
Qy 63 HEATKGIKIDVICVLLQNGAETIRNTDGTALDIA-----DPSAKAVLTGEYK---RDE 114
Db 139 YMAAENHIEVVRFLDNGASGLATEOGFTPLAVALOQGHQDVVSLLENDTKGKVR 198
Qy 115 LLESARSGEEMKMLTPLNVNCHASDGKSTPLHLAAGYNRVKIVOLLQHGADVHAK 174
Db 199 ALHIAARKDDTKAAALLQNDTNADIESKSGFTPLHIAAHYGNINVTLLNRAAVDT 258
Qy 175 DKGDLVPLHNACSYHYEYTELLVYKHGACVAMDLWQFTPLHEAASKRVEVCSLLSYG 234
Db 259 ARNDITPLHVASKRGNNVNNVLLLRDGAIDAKTRDGLTPLHCGARGSGHEQVEMLLDRA 318
Qy 235 ADPTLLNCHNSAIDLAPTLPOLKERLAYEFKGHSLQAAAREADVTRIKKHLSELMVNEKH 294
Db 319 A-PILSKTN-----GLSLPHMATQGD-----HLNCVOLLQOH 350
Qy 295 -----POTETALHCAAAPYPRKQICELLRLKRGANINEKTEFTPLHVASEKAHN 347
Db 351 NVPVDVNTDYLTLHVAACHGCHYK---VAKVLLDKKANPAKALNGFTPLHIAACKNRI 407
Qy 348 DVVEVVVHKAQVNALDNLGQTSLSHRAAYCGHLQTCRLLLSYGCDPNILISQGTALOMG 407
Db 408 RMELLKRGASIQAVTSRGFTPLHVAAFMGHVNITVSLMHGASPNNTNVRGETALHMA 467
Qy 408 -----NENVOQLQBSISLGNSEADQ---LLEAAKAGADVETVKK-LCTVQSVNCRDIEGR 459
Db 468 ARSGQAEVVRVLYVQDGAQVEAKARDEQHTPHFSRLGAEIVQVLYQVQASGPAATTS- 526
Qy 460 QSTPLHFAAGYNRVSVYVYLLQHGADVHAKDKGGLVPLHNACSYGHYEAELLVKGAVV 519
Db 527 -YTPLHLHSAEGHEEDVAFLDHDGAFSLTITTKGFTPLHVAAYKGLVASLLQKAS 585
Qy 520 NVADLWKTPLHEAAAKGKYETICKLLLLQHGADPTKKNRDGNTPDL-L-VKDGDTDQDLL- 577

```

```

Db 586 DAAGKSGLPFLHVAHYDNQKVALLLDQASPHAAAKNGYTPPLHTAAKKNQMDIATSL 645
Qy 578 -----RGDAALLDAAKGGLARVKKLSSPD-NVNCRTQGRHSTPLHLAAGYNN 625
Db 646 EYCADANAVTROGIAVSHLAQEGHVDVYSLLSRANVNLNKRGLN--PLHLGGQEDR 703
Qy 626 LEVAEYVLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFPLHE 685
Db 704 VNVAEVYVLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFPLHE 763
Qy 686 AAQKGTQLCALLAHGADPTLKNQEGTPLDLVS-----ADVYSALLTAAMPSPSALP 738
Db 764 AAQOGHTHIINVLLQNNASPNELTVNGNTALAIARLGLYISVVDTLKVVTEELMTITIT 823
Qy 739 SCYK---POVLNGVRSPGATADALSGSPSSLSAASLDNLSSFSSELSSVSSSGT 795
Db 824 EXHKMNVPEMTNFEVD--MSDDEVGKG-SAPEKLSHGEYI-----SDGEE 865
Qy 796 GASLLEKKEVPGVDFSIQFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIG 850
Db 866 G-----EDAITGDTKSLGPODLKELG 887

RESULT 9
Q13768 PRELIMINARY; PRT; 1719 AA.
AC Q13768;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Alt. ankyrin (variant 2.2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEMATOPOIETIC;
RX MEDLINE=90159830; PubMed=2137557;
RA Lux S.E., John K.M., Bennett V.;
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT structure with homology to tissue-differentiation and cell-cycle
RT control proteins.";
RL Nature 344:36-42(1990).
DR EMBL; X16609; CA34611.1; -.
DR HSP; Q00420; IAWC.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZUS; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 22.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1719 AA; 189011 MW; F63465D16D975CBF CRC64;

Query Match          14.6%; Score 842; DB 4; Length 1719;
Best Local Similarity 29.9%; Pred. No. 7e-45;
Matches 240; Conservative 122; Mismatches 264; Indels 177; Gaps 20;

Qy 3 GRKDVVEYLLQNGASVQARDGGLPLHNACSFHAEVYVNLRLRHGADPNARDNNYTP 62
Db 89 QDVEVYVLLQNGASVQARDGGLPLHNACSFHAEVYVNLRLRHGADPNARDNNYTP 148
Qy 63 HEATKGIKIDVICVLLQNGAETIRNTDGTALDIA-----DPSAKAVLTGEYK 113

```



```

Db 149 AVALQOQHENVVAHLINYGKGVRLPALHIAARNDDTRTA-----AVLQNDPNP 199
Qy 114 ELLESARSGNEKMMALLTPLNVNCHASDGRKS-----TPHLHAAGY 155
Db 200 DVL--SKTG-----FTPLHIAAHYENLVAQLLNRCASVNFTPONGITPLHIASRR 249
Qy 156 NRKIVOLLQHGADYHAKDGDVPLPHNACSXGHEVTELLVVKHGACVNMALDWFOTPL 215
Db 250 GNVIMVRLLLDRGAQIETKTKDELTPURCAARNGHVRISILLDHGAPIAKTKNGLSPI 309
Qy 216 HEARKNRVVCSSLLLSYGADPTLLNCHNKSADLAPTQPKERLAYEFKHSLLQARE 275
Db 310 HMAAQGHDLQCVRLQYDAE-----ID----- 332
Qy 276 ADVTRIKKHLSEMVNEKHPQTHETALHCAASPYPKRKOICELLRKGANINEKTRFEL 335
Db 333 -DIT-----LDHLTPLHVA-----HCCHHVRVAKVLLDKGAKPNSRALNGF 372
Qy 336 TPLHVASEKAHNDVVEVVVHKEAKVNALDNLGQTSLHRAAYCGHLQTCRLLSYCGDPNI 395
Db 373 TPLHIACKKNHVRVMEILLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLORGASPNV 432
Qy 396 ISLOGFTALQW-----NENVQQLQE-----GISLGNSEADROLLE- 432
Db 433 SNVKVETPLHMAARAGHTEVAKYLLQNKAKVNKAKAKDDQTPHCAARIGHTNMVYKLLLEN 492
Qy 433 -----AAKAGDVETVKKLTQV--SYNCRDIEGROSTPLHFAAGYNRVS 474
Db 493 NANPNLATTAGHTPLHIAAREGHVETVALLKEASQACMTKG--FTPLHVAAYKGRV 550
Qy 475 VVEYLLOHGADVHAKDGGVPLPHNACSXGHEVTELLVVKHGAVNVNADLWK-FTPLHEA 533
Db 551 VAEILLERDAHPNAAGKNGLTPLHVAVHNNLDIVKLLPRGGSPPH-SPAWNRYTPLHIA 609
Qy 534 AAKGYEICKLLOHGADPTKKNRDGNTPLDL-VKGDGTDIQLDRLGDAALLDAKKGCL 592
Db 610 AKQNVVARSLLQYGGSSANAESVQGVTPHLHAAQEGHAEVALL----- 654
Qy 593 ARVKKLSPPNVNCRDTQGRHSTPLHIAAGYNNLEVAEYLLQHGADVNAODKGGILPHN 652
Db 655 -----LSQANGNLNKGSG--LTPHLHVAQEGHPVADVLIKGVWVDATTRMCTYPLHV 707
Qy 653 AASYGHVDVVAALLIKYNACVNATDKWAFTPLHAAQKRGRTQLCALLAHAGADPTLKNOEG 712
Db 708 ASHYGNIKLVKFLQHOADVNAKTKGYSPLHQAQOQGHDTIVTLLKNGASPNVSSDG 767
Qy 713 QTPLDL-----VSADDVSALLT 729
Db 768 TTPHIAAKRLGYISVTDVLKVV 790

RESULT 10
P97582
ID P97582 PRELIMINARY; PRT; 843 AA.
AC P97582;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Ankyrin (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=BRAIN;
RA Cosentino M.T., Jones O.T.;
RT "Rat brain ankyrin membrane binding domain."
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U65916; ABA47551.1;
DR HSSP; Q00420; IAWC.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 24.

```

```

DR SMART; SM00248; ANK; 21.
DR PROSITE; PS00088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW ANK repeat; Repeat.
FT NON_TER 843 843
SQ SEQUENCE 843 AA; 89982 MW; DC5A6AC78825D223 CRC64;

Query Match 14.6%; Score 841.5; DB 11; Length 843;
Best Local Similarity 29.7%; Pred. No. 2.6e-45;
Matches 265; Conservative 138; Mismatches 351; Indels 137; Gaps 21;

Qy 3 GRKDVVEYLLQNGASVQARDGGLPLHNACSFHGAHYVNNLLRHGADPNARDNMNYTPL 62
Db 40 GHVGLVQLLELGRGSSVSATKKGNTALHIASLAGQAEVVKLVKEGANINAQSQNGFTPL 99
Qy 63 HEAAIKKIDVICIVILLQGAEPTRINDGRT-----ALDADPSAKAVLTGEYKDEL--- 115
Db 100 YMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQOQHNOAVAILLENDTKGKVRP 159
Qy 116 -LESARSGNEEKMMALLTPLNVNCHASDGRKS-----TPHLHAAAGYNRVKIVQL 163
Db 160 ALHTAARKDDTKSAALLQ---NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATL 216
Qy 164 LLOHGADVHAKDKDGLVPLHNACSXGHEVTELLVVKHGACVNMALDWFOTPLHEAASKNR 223
Db 217 LLNRGAADVFTARNGITPLHVASKRGNTNMVKLLLDRCGQIDAKTRDGLTPLHCAARSGH 276
Qy 224 VECSLLLSYGADPTLLNCHNKSADLAPTQPKERLAYEFKHSLLQAAAREADVTRKK 283
Db 277 DOVVVELLIERGA-PLLARTKN-----GLSPLHMAAQGDHVECVK 314
Qy 284 HLSLEMVNFKHP-----QTHETALHCAASPYPKRKOICELLRKGANINEKTRFELPL 338
Db 315 HL-----LOHKAPVDVTDLYTLALHVA--HCGHYRVTKLLLDKRAPNARALNGFTPL 367
Qy 339 HVASEKAHNDVVEVVVHKEAKVNALDNLGQTSLHRAAYCGHLQTCRLLSYCGDPNIISL 398
Db 368 HIACKKNRIKVMELLVKYGAVIAITESGLTPIPVAAFMGHLNIVLLLLONGASPDVTNI 427
Qy 399 QGFTALQMGNEVQQLQEGISLGNSEADROLLEAAKAGDVETVKKLTQVSVNCRDIEG 458
Db 428 RGETALHN-----AARAGEVEVWR--CLLRNGALVDARA 459
Qy 459 R-QSTPLHFAAGYNRVSVEVYLLQHGADVHAKDGGVPLHNACSXGHEVVAELLVKHGA 517
Db 460 REEQTPHLIASRLGKTEIVQLLQHMHPDAATTNGYTPHLISAREGQVDVASVLEAGA 519
Qy 518 VVNVADLWKFTPLHEAAAKGYEICKLLOHGADPTKKNRDGNTPLDLVKDGDTDIQLDL 577
Db 520 AHSLATKKGFTPLHVAAYKGSLLDVAKLLQRRRAADSAGKNGLTPLHVAAHYDN----- 573
Qy 578 RGDAAALLDAKKGCLARVKKLSSPDVNVNCRDTQGRHSTPLHIAAGYNNLEVAEYLLQHGA 637
Db 574 -----OKVALLLEKGAEP-----HATAKNGYTPHIAAKKNQMOQTASTLLNTGA 618
Qy 638 DVNAQDKGGLTPLHNAASGHHVDVAALLIKYNACVNATDKWAFTPLHAAQKRGRTQLCAL 697
Db 619 ETNIVTKGVVTPHLASQEGHTDWTVLVEKGANIHMTKSGLTSLHAAEEDKYNVADI 678
Qy 698 LLAHGADPTLKNOEGQTPLDLV-----SADDYSALLTAAMPSPA-LPSCYKP----- 743
Db 679 LTKHGADQDAYTKLGYTPLIVACHYGNVKNMNVNLLKOGANVNAKTKNGYTPHQAQOQH 738
Qy 744 -QVLNGVRSPGATADALSSGSPSSLSAASLDNLGSFSELSVSVSSSGTEGASSLEK 802
Db 739 THIIINVLLQHGAKPNATTA--NGNTALAIKRLGYI--SVVDTLKVVTEEVTTTTTITE 794
Qy 803 KEVPGVDFSLITQFVRNLGLEHLMIDIFERE--QITLDVLVEMGHKELKEIG 850
Db 795 KHLNAPETMTE-----VLDVSDDEGDDTVTGDGGEYLRPBDLRELK 836

```



```

101-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Erythroid ankyrin.
GN ANK1 OR ANK-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=NEURAL;
RX MEDLINE=93252825; PubMed=9486643;
RA Birkenmeier C.S.; White R.A.; Peters L.L.; Hall E.J.; Lux S.E.;
RA Barker J.E.;
RT *Complex patterns of sequence variation and multiple 5' and 3' ends
RT are found among transcripts of the erythroid ankyrin gene.";
RL J. Biol. Chem. 268:9533-9540(1993).
DR EMBL; X69063; CAA48801.1; -.
DR HSSP; Q00420; IAWC.
DR MGD; MGI:88024; ANK1.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 22.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1848 AA; 202577 MW; 5744BECBF9EBA056 CRC64;
Query Match 14.6%; Score 839.5; DB 11; Length 1848;
Best Local Similarity 29.5%; Pred. No. 1.1e-44;
Matches 237; Conservative 127; Mismatches 262; Indels 177; Gaps 20;
QY 3 GRKDVEYLLONGASVOARDGGIPLHNACSFQGHAEVNVNLLRHGADPNARDNNWYTPL 62
DB 93 QGDEVRELVYGANVNAQSOGTPTLYMAAQNHEVVKFLLGANGONVATEDGFTPL 152
QY 63 HEAAIKGKIDYCVILLQHGAEFTI-----RNTDGRATDLADPSAKAVLTGEYKDD 113
DB 153 AVALQOGHENVVAHLINGYTKGVRLPALHIAARNDDRTA-----AVLLONDNP 203
QY 114 ELLESARGNEEKMMALLTPLNVNCHASDGRKS-----TPHLAAGY 155
DB 204 DVL--SKTG-----FTPLHIAAHYENLVNAQILLNRGASVNPFPONGITPLHIASRR 253
QY 156 NRKVIQVLLQHGADVHAKDKGDIPLHNACSYGHYEVTTELLVKGACVNMDLWQFTPL 215
DB 254 GNVIWVRLLDGAGIETRTKDEITPLHCAARNHVRISLEILLHGAPIQAKTNGLSPI 313
QY 216 HEAAKSRVYCVSLLSYGADPTLLNCHNKSATDLAPTPOLKERLAYEFKGHSLQAARE 275
DB 314 HMAAQGHDCVRLLLQYNAE-----ID----- 336
QY 276 ADVTRIKHLSLEMVNFKHPOTHETALHCAASYPKPKQICEILLRKGANINEKTEFL 335
DB 337 -DIT-----LDHLTPLHVA--HCGHHRVAKVLLDKGAPNSKALNGF 376
QY 336 TPLHVASKAHNDVVEVVVYKHEAKVNALDNLGQTSLHRAACVGHLOTCRILLSYGCDPNI 395
DB 377 TPLHIACKNHNIRWELLTKTGASIDAVTESGLTPLHVASFPMGLPIVKNLLQKGASNP 436
QY 396 ISLQCFATQM-----GNENVQQLQEGISLGNSEAD-----ROLLE- 432
DB 437 SNRVETPLHMAARAGHTEVAKYLLONKAKANAKAKDDQTPHCAARTIGHTGMVKLLLEN 496
QY 433 -----AAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVS 474
DB 496 -----

```

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*";
RL Science 287:2185-2195(2000).
DR EMBL: AE003845; AAF59369.2; -.
DR HSSP: P80144; 2MYO.
DR FlyBase: FBgn0011747; Ank.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000488; Death.
DR InterPro: IPR000906; ZU5.
DR Pfam: PF00023; ank; 23.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; ZU5; 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 21.
DR SMART: SM00005; DEATH; 1.
DR SMART: SM00218; ZU5; 1.
DR PROSITE: PS50088; ANK_REPEAT; 21.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KW ANK repeat: Repeat.
SQ SEQUENCE 1549 AA; 170167 MW; 3ECE852B1484DFF CRC64;

Query Match 14.5%; Score 836; DB 5; Length 1549;
Best Local Similarity 27.4%; Pred. No. 1.4e-44;
Matches 298; Conservative 166; Mismatches 469; Indels 154; Gaps 32;

QY 3 GRKDVEYLLONGASVOARDGGLIPLHNACSGHAEVNVLLRHGADPNARDNNWYTPL 62
DB 84 GYDIDCELLRRGIKIDNATKKGNTALHTIASLAGQHDVINQLILYNNANVQSLNGFTPL 143

QY 63 HEAAIKGIDVCIVLLQHGAEPTIRNTDGRITDLDLA-----DPSAKAVLTGEYK---KDE 114
DB 144 YMAAQENHDMCCRTLLANGANSISTEDGFTPLAVAMQOQHDKIVAVLLENDVRGKVRLP 203

QY 115 LLESARGNEKMMALITPLNVNCHASDGRKSTPLHLAGYNRVKIVQLLQHGADVHAK 174
DB 204 ALHIAAKNDVNAKLLQLDHPNADIVSKSGFTPLHIAAHYGNVDIATLLNNKADVNVY 263

QY 175 DKGDVPLHNACSYGHEYTELLVKVIGACVNMADLWQFPLHEAASKNRVEVCSLLISY 234
DB 264 AKHNITPLHVACKMGKLSLCTLLCRGAKIDAAATRDGLTPLHCASRSGHVEYIKHLQQN 323

QY 235 ADPTLLNCHN-KSAIDLAPTQPKRLAYEFKGHSLLOQAAREADVTRIKKHSLE----- 288
DB 324 A-PILFTKNGLSALHMAAQE-----HDEAAHLLLDNKPVDVTVLYLTALHVAHC 376

QY 289 -----MVNFK-HPQTHE-----TALHCAASPYPKRKOICELLRKGANINEKTFEL 335
DB 377 GHVKVAKLLLDYKANPNARALNGFTPLHIAACK--NRIMKVELLIKHGANIGATTESGL 433

QY 336 TPLHVASERAHNDVVEVVKHKAVALDNIGTSLHRAAYCGHLOTCLLLSYGCDPNI 395
DB 434 TPLHVASFGMCINVIYLLQHEASADLPTTRGFTPLHAAANQADITIRILLR-SAKVDA 492

QY 396 ISLQGFAL----QMGNEN-VQQLQEGISLGNSEADR--QLLEAAKAGDVETVKKLCTV 448
DB 493 IAREGOTPLHVASRLGNINIIMLLQHGAFAINAQSDNKYSALHIAKEGENIVQVLLN 552

QY 449 QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHEY 508
DB 553 GAEN-NAVTKKGFTPLHACKYQKQNVQVILLQNGASIDFQGNVDVTPHLHVATHYNPSI 611

QY 509 AELLVKHGAVVNVADLWKTTPPLHEAAKGYEICKLLQLQHGADPTKKNRDGNTPLDL-VK 567
DB 612 VELLKNGSPNICARGOQCAIHCCKKNYLETAMQLQHGADVNTISKSGFPLHAAQ 671

QY 568 DGDITDQLLRGDAALLDAAKGCLARVKKLSLSPDNNVNCRDITQGRHSTPLHLAGYNNLE 627
DB 672 GGNVDVQVLLLEGVISAAKNGL-----TPLHVAAQEGHVL 708

QY 628 VAEYLLQHGADVNAQDKGGIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFTPLHEAA 687
DB 709 VSQILLIEHGANISERTNGYTPPLHMAHYGHLDLVKPFIEADADIEMSNIGTYTPLHQA 768

QY 688 QKRTQCALLLAHGADPTLKNQEGQTPDLVSDADDVSALLTAAMPSPALPSYCKKQVNLN 747
DB 769 QOQHIMIINLLRHKANPNALTDKGNTALHASN-----LGVYVTMESLKIVTSTSVIN 822

QY 748 G-----VRSPGATADALSGSPSSLSAAS-----SLDNLGSFSELSVWS 790
DB 823 SNITCAIEELKLVMTPELMQETLLSDSDDECDLLDHNHYKYMATDDLKANYGQDQKNFD 882

QY 791 SSQTE-----GASSLEKKEVPGVDFSIQFVRNLGLEHMD--IFERQITLDLVEMGHK 844
DB 883 TTNTDHLDTDSVLNKKIEILPNEMSCIELTE---IGHKPDNVVIARSQVHLGFLVSP--- 936

QY 845 ELKEIGINAYGHRHKLKIGVE-----RLISGOQGLNPYLTLTSSGTLID 891
DB 937 LVDARGGSMRGYRHNGVRIIVPPKACAEPTRTICRVYKQPVVNPPLMGEALVSRILE 996

QY 892 LSPDDKEFOS---VEEMQSTVREHRDGGHAGGIFNRYNLIKQKVCNKKLWERYTHRRK 948
DB 997 MSPVDGMLFSPITLEVPHYGTLRK-----NEREII-LRSDNGESWREH-NLYK 1043

QY 949 EYSEENHNANERMLFHGSPFYNAIIHKGFDERHAYIGMGFAGIYPAENSSKSNQVYVG 1008
DB 1044 DIIGEDINQTEE---FHSDDRIVR-IVTQNVPHFFAVV-----SRVQREVHV 1085

QY 1009 IG--GGT 1013
DB 1086 IGPDDGT 1092

RESULT 14
Q24241 PRELIMINARY; PRT; 1549 AA.
AC Q24241;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE ANKYRIN
GN ANK OR CG1651.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A631;
RX MEDLINE=95024098; PubMed=7937942;
RA Dubreuil R.R., Yu J.;
RT "Ankyrin and beta-spectrin accumulate independently of alpha-spectrin
in *Drosophila*.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:10285-10289(1994).
DR EMBL: L35601; AAC37208.1; -.
DR HSSP: P80144; 2MYO.
DR FlyBase: FBgn0011747; Ank.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000488; Death.
DR InterPro: IPR000906; ZU5.
DR Pfam: PF00023; ank; 23.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; ZU5; 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 21.
DR SMART: SM00005; DEATH; 1.
DR SMART: SM00218; ZU5; 1.
DR PROSITE: PS50088; ANK_REPEAT; 21.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KW ANK repeat: Repeat.
SQ SEQUENCE 1549 AA; 170096 MW; 587FB6BF88B657CE CRC64;

Query Match 14.5%; Score 834; DB 5; Length 1549;
Best Local Similarity 24.8%; Pred. No. 1.9e-44;
Matches 288; Conservative 175; Mismatches 397; Indels 300; Gaps 32;

QY 3 GRKDVEYLLQNGASVQARDGGILPLHNACSFHAEVYVNNLLRHGADPNARDNNWYTPL 62
DB 84 GYVDICCELLRRGKIDNATKNGTALHIASLAGQHDYINQLILYNANVNVQSLNGETPL 143
QY 63 HEAAIKGKIDVICVILLOHCAEPTIRNTDGRALDADPSAKAVLTGEYKKDELLESARG 122
DB 144 YMAOENHDNCRTLLANGANSLESTEDGFT-----PLAVAMQOG----- 183
QY 123 NEEKMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLQHGADVHAKDKGLVP 181
DB 184 -HDKIVAVLLENDVR-----GKVRPLPALHIAAKNDVNAKLLLOHDPNADIVSKSGFTP 237
QY 182 LHNACSYGHYEVELLVKHGACVNAQMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241
DB 238 LHIAAHYGNVDIATLLNNKADVNTVAKHNITPLHVACKWGKLSLCTLLLCRGA----- 291
QY 242 CHNKSADLAPTPOLKERLAYEFKHSLLQAREADVTRIKKHSLEMVNFKHPQTHETA 301
DB 292 -----KIDAATRDGL-----TP 303
QY 302 LHCAASPYPKKQICELLRRKGANINEKTEFLPLHVASEKAHNDVVEVVKHEAKVN 361
DB 304 LHCASRSGH---VEVILKHLQQNAQPIPLTKNGLSALHWAAGOEDEAAHLLLDNKAQPD 360
QY 362 ALDNLGQTSLHRAAYCGHLLQTCRLLSYGCDDNITISLOGFTALONGNEN-----VOQLLQ 416
DB 361 ETVTDYLTALHVAHCGHVKAQKLLDYKANPNARALNGFTPLHIAACKNRKIKVVELLIK 420
QY 417 EGISLG-----NSEADROLLE-----AAKAGDVEF 441
DB 421 HGANIGATTESGLTPLHVASFNGCINVIYLLQHEASADLPTIRGETPLHIAARANQADI 480
QY 442 VRKLCITQSVNCRDIEGROSTPLHFAAGYNRVSVVEYVILLOHGADVHAKD----- 490
DB 481 IRIILRSKAVDAIVREGO--TPLHVASRLGNIIMILLQHGAEINAQSNOKYSALHIAA 538
QY 491 -----KGLVPLHNACSYGHVEAEVALLVKGAVNVNADLWKF 528
DB 539 KEGQENIVQVLENGAENNAVTCKGFTPLHLACKYQKQNVVQILQNGASIDFCQKNDVT 598
QY 529 PLHEA-----AAKGKY-EICKLILLOHGADPTKK 555
DB 599 PLHVATHYNNPSIVELLLKNGSSPNLCARNGOCATHIACKKNYLEIAMOLLOHGADVNI 658
QY 556 NRDGNTPLDL-VKDGDTDITQDILLRGAALLDAAKKGCLARVKKLSSPNDVNCRDTOGRHS 614
DB 659 SKSGFSPHLAAQGGNVDMVQLLLEYGVISAAAKNGL----- 695
QY 615 TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNACSYGHVDVVAALLIKYNACVNA 674
DB 696 TPLHVAQEGHVLVQSILLLEGANISERTFRNGYTPPLHMAHYGHLDLVKFFTEINDADIE 755
QY 675 TDKWAFTPLHEAAOKRTOICALLAHGADPTLKNOEGOTPLDVSADVSALLTAAMP 734
DB 756 SSNIGYTPHLQAQOQHIMIINLLRHKANPNALTKDGN TALHASN-----LGIVTVM 809
QY 735 SALPSCYKQVQLNG-----VRSPGATADALSSGSPSPSSLSAAS-----SLDN 777
DB 810 ESLKIVTSTSVINSNIGATEEKLKVMYTPELMQETLLSDSDDECDLDDHNNHYKMYMATDD 869
QY 778 LSGSPSELSSVSSSGTE-----GASLEKKEVPDGVDFSIQVRLNGLEHLM-----IFERE 831
DB 870 LKANYQDQKNEDTTNTDHLTDVSVINKKEILPNEMSCIELTE-----IGHKPDNVVIARS 926
QY 832 QITLDVLVEMGHKELKEICINAYGHRKLIKGE-----RLISQOQGLNYPYL 878
DB 927 QVHLGLVLSF---LYDARGSGRMGRYHNGVRIIVPPKACAEPTTRITCRYVKVQPVVNPP 983

QY 879 TLANTSGGTILIDSLPDDKEFQS---VEEQMQSTVREHRDGGHAGGIFNRYNLIKIQKVC 935
DB 984 LMEGEALVSRIEMSPVDGMFLSPITLEVPHYGTLRK-----NEREII-ILRSD 1031
QY 936 NKKLWEYTHRRKEVSEENINHANERMLFHSQSPFVNAIIHKGFDERHAYIGMGAGLIYF 995
DB 1032 NGESWREH-NLYKDITGEDINQTEE---FHSDRIVR-IVTONVPHFFAVV----- 1076
QY 996 AENSXSKSNQVYVIGG--GGT 1013
DB 1077 ----SRVRQEVHVIGPDGGT 1092

RESULT 15
Q9NCP8 PRELLIMINARY; PRT; 1159 AA.
ID Q9NCP8
AC Q9NCP8; 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Ankyrin 2.
GN ANK2 OR CG7462.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20304926; PubMed=10844021;
RA Bouley M., Tian M.-Z., Paisley K., Shen Y.-C., Malhotra J.D.,
RA Hortsch M.;
RT "The Ll-type cell adhesion molecule neuroglian influences the
RT stability of neural ankyrin in the Drosophila embryo but not its
RT axonal localization.";
RL J. Neurosci. 20:4515-4523(2000).
DR EMBL; AF190635; AAF73309.1; -;
DR HSSP; P42773; 11HB.
DR FlyBase; FBgn0017645; Ank2.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00791; ZU5; 1.
DR SMART; SM00248; ANK; 21.
DR SMART; SM02118; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR KW ANK repeat; Repeat.
SQ SEQUENCE 1159 AA; 125769 MW; 9ED146C9E483DE58 CRC64;

Query Match 14.2%; Score 820; DB 5; Length 1159;
Best Local Similarity 27.9%; Pred. No. 9.9e-44;
Matches 268; Conservative 138; Mismatches 373; Indels 180; Gaps 21;

QY 3 GRKDVEYLLQNGASVQARDGGILPLHNACSFHAEVYVNNLLRHGADPNARDNNWYTPL 62
DB 55 GHIVVSELLRGAIVDSATKKGTALHIASLAGOEVEVVKLLHNASVNVQSONGFTPL 114
QY 63 HEAAIKGKIDVICVILLOHCAEPTIRNTDGRALDADPSAKAVLTGEYKKDELLESARG 122
DB 115 YMAOENHDVAVRLLLSNGANOSLATEGFT-----PLAVAMQOG----- 154
QY 123 NEEKMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLQHGADVHAKDKGLVP 181
DB 155 -HDKVAVLLESDTR-----GKVRPLPALHIAAKKDDVKAATLLLDNDHNPVTSKSGFTP 208
QY 182 LHNACSYGHYEVELLVKHGACVNAQMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241
DB 209 LHIAASHYGNQNIANILLQKQADVNSAKHNISPLHVAAGKGTNNVSLLEKGG-----N 263
QY 242 CHNKSADLAPTPOLKERLAYEFKHSLLQAREADVTRIKKHSLEMVNFKHPQTHETA 301
DB 264 TEAKTRDGLTP----- 274

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:31:44 ; Search time 42.3932 Seconds
(without alignments)
3897.581 Million cell updates/sec

Title: US-09-843-159B-4
Perfect score: 6464
Sequence: 1 RCSARRGAAGGAGGARGARV.....AYPEVLITYQIMRPGMVDG 1240

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	6464	100.0	1240	22	AA97749
2	6375.5	98.6	1262	22	AA97749
3	6375.5	98.6	1385	22	AA97749
4	6121	94.7	1169	22	AA97749
5	6121	94.7	1169	22	AA97749
6	6108	94.5	1166	22	AA97749
7	6092	94.2	1166	22	AA97749
8	6053	93.6	1166	22	AA97749
9	5766	89.2	1100	22	AA97749
10	5605	86.7	1074	20	AA97749

11	5543	85.8	1166	22	AA97749	Human SPANK. Homo
12	5474	84.7	1431	21	AA97749	Human tankyrase II
13	5103.5	79.0	1327	21	AA97749	Human tankyrase I
14	5103.5	79.0	1327	21	AA97749	Human tankyrase
15	5103.5	79.0	1327	22	AA97749	Human tankyrase1 S
16	5103.5	79.0	1327	23	AA97749	Human tankyrase
17	5040.5	78.0	1333	21	AA97749	Human tankyrase II
18	5038.5	77.9	991	22	AA97749	Mouse SPANK. Mus
19	4624.5	71.5	907	22	AA97749	Human breast cancer
20	4102	63.5	784	22	AA97749	Human tankyrase2 c
21	4005	62.0	1181	22	AA97749	Drosophila melanog
22	3997	61.8	1181	22	AA97749	Drosophila tankyrra
23	3959	61.2	756	22	AA97749	Human tankyrase2 c
24	3938	60.9	802	23	AA97749	Protein of APP rel
25	3307	51.2	949	21	AA97749	Human truncated ta
26	3307	51.2	949	23	AA97749	Truncated human ta
27	2722	42.1	522	22	AA97749	Human tankyrase2 c
28	2165	33.5	415	22	AA97749	Human Grb7 effecto
29	2132.5	33.0	673	21	AA97749	Human truncated ta
30	2132.5	33.0	673	23	AA97749	Truncated human ta
31	2056.5	31.8	1099	22	AA97749	Human tankyrase2 e
32	1942	30.0	385	22	AA97749	Human T cell leuka
33	1895	29.3	362	22	AA97749	Human TRF1 TANK2 b
34	1706	26.4	331	22	AA97749	Human breast cancer
35	1395	21.6	261	22	AA97749	Human protein sequ
36	1251	19.4	240	21	AA97749	Human ORF3133
37	892.5	13.8	4274	22	AA97749	Novel human diagno
38	892.5	13.8	4386	22	AA97749	Novel human diagno
39	886.5	13.7	215	22	AA97749	Human tankyrase2 c
40	877	13.6	1872	22	AA97749	Human protein SEQ
41	874	13.5	160	21	AA97749	Human tankyrase II
42	874	13.5	160	22	AA97749	Human tankyrase I
43	869.5	13.5	4397	22	AA97749	Novel human diagno
44	868.5	13.4	1549	22	AA97749	Drosophila melanog
45	868.5	13.4	1549	22	AA97749	Drosophila melanog

ALIGNMENTS

RESULT 1
AA97749
ID AA97749 standard; Protein; 1240 AA.
XX
AC AA97749;
XX
XX
DT 06-AUG-2001 (first entry)
XX
XX Tankyrase homologue isotype 2 protein sequence.
DE
DE Tankyrase homologue isotype 2; TaHo-1; TaHo-2; cell proliferation;
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine.
XX
XX Unidentified.
XX
XX WO200130987-A2.
XX
XX 03-MAY-2001.
XX
XX 25-OCT-2000; 2000WO-US41528.
XX
XX 25-OCT-1999; 99US-0427154.
XX
XX (RIGI-) RIGEL PHARM INC.
XX
XX Luo Y, Chan E, Xu X, Huang B;
XX WPI: 2001-300503/31.
XX N-PSDB; AA91488.
XX
XX Novel recombinant cell cycle polypeptide, tankyrase H useful for
PT inducing or preventing cell proliferation in cells, and for diagnosing,
PT

PS Example 2; Page 173-176; 242pp; English.

CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.

XX SQ Sequence 1262 AA;

Query Match 98.6%; Score 6375.5; DB 22; Length 1262;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1225; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 6 RGAAGGCGAQRGARGAAGTAPDPVTAGSQAARALSSAPGGALLLAGGLLLRLAL 65
DB 29 RGSRGAGSPARGAR-GRGHGTAPDPVTAGSQAARALSSAPGGALLLAGGLLLRLAL 87
QY 66 LLAVAAARIMSGRCAGGACASAAAEVPAARELFACRNGDVERKRLVTPKVN 125
DB 88 LLAVAAARIMSGRCAGGACASAAAEVPAARELFACRNGDVERKRLVTPKVN 147
QY 126 RDTAGRKSTPLHFAAGFGKRDVVYLLQNGANVQARDGGGLPLHNACSGFHAEWNLL 185
DB 148 RDTAGRKSTPLHFAAGFGKRDVVYLLQNGANVQARDGGGLPLHNACSGFHAEWNLL 207
QY 186 RHGADPNARDNWTPLHEAAIKGIDVCIIVLLQHGAEPTIRNTDGRALDADPSAKAV 245
DB 208 RHGADPNARDNWTPLHEAAIKGIDVCIIVLLQHGAEPTIRNTDGRALDADPSAKAV 267
QY 246 LTGEYKDELLSARSNEEKMMALLPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLL 305
DB 268 LTGEYKDELLSARSNEEKMMALLPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLL 327
QY 306 QHGADVHAKDKGLVPLHNACSYGHEVTELLVKGACVNAIDLWQFTPLHEAASKNRVE 365
DB 328 QHGADVHAKDKGLVPLHNACSYGHEVTELLVKGACVNAIDLWQFTPLHEAASKNRVE 387
QY 366 VCSLLLSYGADPTLLNCHNKSADLAPTQPKERLAYEFKSHLSLQAAAREADVTRIKKHL 425
DB 388 VCSLLLSYGADPTLLNCHNKSADLAPTQPKERLAYEFKSHLSLQAAAREADVTRIKKHL 447
QY 426 SLEMVNFKHPOTHEALHCAASAPYPRKQICEILLRKGANINEKTEFTPLHVASEKA 485
DB 448 SLEMVNFKHPOTHEALHCAASAPYPRKQICEILLRKGANINEKTEFTPLHVASEKA 507
QY 486 HNDVVEVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDDPNIISLQGTALQ 545
DB 508 HNDVVEVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDDPNIISLQGTALQ 567
QY 546 MGENVOQLLOEGLISLNSENDROLLBAKAGDVETVKLCITVQSVNCRDIEGRQSTPLH 605
DB 568 MGENVOQLLOEGLISLNSENDROLLBAKAGDVETVKLCITVQSVNCRDIEGRQSTPLH 627
QY 606 FAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNACSYGHEVTELLVKGACVNAIDLW 665
DB 628 FAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNACSYGHEVTELLVKGACVNAIDLW 687
QY 666 KFTPLHEAAAKGVEICKLLQHGADPTKKNRDGNTPDLVKGDDTDITQDLLRGDAALLD 725
DB 688 KFTPLHEAAAKGVEICKLLQHGADPTKKNRDGNTPDLVKGDDTDITQDLLRGDAALLD 747
QY 726 AAKKGCLARVKKLSPPNVNCRDTQGRHSTPLHLAAGYNLNLEVAEYLLQHGADVNAQDKG 785
DB 748 AAKKGCLARVKKLSPPNVNCRDTQGRHSTPLHLAAGYNLNLEVAEYLLQHGADVNAQDKG 807
QY 786 GLTPLHNAASGVHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTOCALLHAHADP 845
DB 808 GLTPLHNAASGVHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTOCALLHAHADP 867

QY 846 TLKNOEGOTPLDLVSADDVSALLTAAMPSPALPSCYKFOVLNGVRSPGATADALSSGPSS 905
DB 868 TLKNOEGOTPLDLVSADDVSALLTAAMPSPALPSCYKFOVLNGVRSPGATADALSSGPSS 927
QY 906 PSSLSAASSLNLGSGFSFSELSSVSSSGTEGASSLEKKEVPGVDFTITQFVRNLGLEHLM 965
DB 928 PSSLSAASSLNLGSGFSFSELSSVSSSGTEGASSLEKKEVPGVDFTITQFVRNLGLEHLM 987
QY 966 DIFEREQITLDLVEMGHKELKEICINAYGHRHKLKIGVERLISGOQGLNPYLLTNTSGS 1025
DB 988 DIFEREQITLDLVEMGHKELKEICINAYGHRHKLKIGVERLISGOQGLNPYLLTNTSGS 1047
QY 1026 GTILIDLSPPDKKEFQSVSEEMQSTVREHRDGGHAGGIENRYNLIKIQVCNKKLWERYTH 1085
DB 1048 GTILIDLSPPDKKEFQSVSEEMQSTVREHRDGGHAGGIENRYNLIKIQVCNKKLWERYTH 1107
QY 1086 RRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDBRHAYIGGMFGAGIYFAENSSKSNQY 1145
DB 1108 RRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDBRHAYIGGMFGAGIYFAENSSKSNQY 1167
QY 1146 VYIGGGTGCPVHKDRSCYICHROLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSV 1205
DB 1168 VYIGGGTGCPVHKDRSCYICHROLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSV 1227
QY 1206 NGLALAEVVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
DB 1228 NGLALAEVVIYRGEQAYPEYLITYQIMRPEGMVDG 1262

RESULT 3
AAB66294
ID AAB66294 standard; Protein: 1385 AA.
XX AAB66294;
XX 05-APR-2001 (first entry)
XX Human tankyrase2 TANK2-LONG SEQ ID NO: 133.
DE Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW Inflammatory disorder.
OS Homo sapiens.
XX WO200100849-A1.
XX 04-JAN-2001.
XX 28-JUN-2000; 2000WO-US17827.
XX 29-JUN-1999; 99US-0141582.
XX (ICOS-) ICOS CORP.
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
WPI: 2001-102896/11.
N-PSDB: AAF63952.
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX Claim 2; Page 191-194; 242pp; English.
XX The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.

0:

QY 132 KSTPLHFAAGFRKDVVEYLLQNGANVQARDGGLIPLHNACSFHAEVNVNLLLRHGADP 191
DB 61 KSTPLHFAAGFRKDVVEYLLQNGANVQARDGGLIPLHNACSFHAEVNVNLLLRHGADP 120
QY 192 NARDNNWYTPHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYK 251
DB 121 NARDNNWYTPHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYK 180
QY 252 KDELLESARSNEEKMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADV 311
DB 181 KDELLESARSNEEKMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADV 240
QY 312 HAKDKGDLVPLHNACSYGHEVTELLVKGACVNMDLWQFTPLHEAASKNRVEVCSLLL 371
DB 241 HAKDKGDLVPLHNACSYGHEVTELLVKGACVNMDLWQFTPLHEAASKNRVEVCSLLL 300
QY 372 SYGADPTLLCHNKSAIDLAFTPOLKERLAYEFKGHSLQAAAREADVTRIKKHLSELVN 431
DB 301 SYGADPTLLCHNKSAIDLAFTPOLKERLAYEFKGHSLQAAAREADVTRIKKHLSELVN 360
QY 432 FKHPQTHETALHCAASAPYKPKQICELLRKGANINEKTKFELTPLHVASEKAHNDVVE 491
DB 361 FKHPQTHETALHCAASAPYKPKQICELLRKGANINEKTKFELTPLHVASEKAHNDVVE 420
QY 492 VVYKHEAKVNALNLGOTSLSHRAAYCGHLQTCRLLLSYGCDPNIISLQGFALQMGNEV 551
DB 421 VVYKHEAKVNALNLGOTSLSHRAAYCGHLQTCRLLLSYGCDPNIISLQGFALQMGNEV 480
QY 552 QOLLOEGISLGNSEADROLLEAAKAGDVETVKIKCTVQSVNCRDIEGRQSTPLHFAAGYN 611
DB 481 QOLLOEGISLGNSEADROLLEAAKAGDVETVKIKCTVQSVNCRDIEGRQSTPLHFAAGYN 540
QY 612 RVSVEYLLQHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKGAVNVADLWKFTPLH 671
DB 541 RVSVEYLLQHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKGAVNVADLWKFTPLH 600
QY 672 EAAKAGYEICKLLQHGADPTKKNRDNTPDLVKGDDTDIODLLRGDAALLDAARKGC 731
DB 601 EAAKAGYEICKLLQHGADPTKKNRDNTPDLVKGDDTDIODLLRGDAALLDAARKGC 660
QY 732 LARVKLSSPDNVCNCRDTCGRHSTPLHLAAGYNVLEVAEYLLQHGADVNAQDKGGLIPLH 791
DB 661 LARVKLSSPDNVCNCRDTCGRHSTPLHLAAGYNVLEVAEYLLQHGADVNAQDKGGLIPLH 720
QY 792 NAASGYHVDVAALLIKYNACVNATDKWFTPLHEAAQKGRQTQCALLLAHGAADPTLKNQE 851
DB 721 NAASGYHVDVAALLIKYNACVNATDKWFTPLHEAAQKGRQTQCALLLAHGAADPTLKNQE 780
QY 852 GQTPDLVVSADVSALLTAAMPSPALPSCYKQVNLGVRSFGATADALSSGSPSSPSLSA 911
DB 781 GQTPDLVVSADVSALLTAAMPSPALPSCYKQVNLGVRSFGATADALSSGSPSSPSLSA 840
QY 912 ASSLDNLGSSFSSELSSVSSSGTEGASSLEKKEVPVDFSIOTFVRNLGLHEHLMDFERE 971
DB 841 ASSLDNLGSSFSSELSSVSSSGTEGASSLEKKEVPVDFSIOTFVRNLGLHEHLMDFERE 900
QY 972 QITLDVLVEMGHKELKEIGINAYGHRHKLKGVRELISGGOGLNPYLTLMTSGSGTILID 1031
DB 901 QITLDVLVEMGHKELKEIGINAYGHRHKLKGVRELISGGOGLNPYLTLMTSGSGTILID 960
QY 1032 LSPDDKEFQSVSEEMQSTVREHRDGGHAGGIFNRYNLTIKQVCNKKLWERYTHRRKEVS 1091
DB 961 LSPDDKEFQSVSEEMQSTVREHRDGGHAGGIFNRYNLTIKQVCNKKLWERYTHRRKEVS 1020
QY 1092 EENHNHANERMLPHGSPFNVAIITHKGFDERHAYIGMFGAGIYFAENSSKNQVYVIGG 1151
DB 1021 EENHNHANERMLPHGSPFNVAIITHKGFDERHAYIGMFGAGIYFAENSSKNQVYVIGG 1080
QY 1152 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHSVTCRPSVNGLALA 1211
DB 1081 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHSVTCRPSVNGLALA 1140

QY 1212 EYVYVRGQAYPEYLITYQIMRPEGVDG 1240
DB 1141 EYVYVRGQAYPEYLITYQIMRPEGVDG 1169

RESULT 5

AAB66288
ID AAB66288 standard; Protein; 1169 AA.

XX AAB66288;

XX AC 05-APR-2001 (first entry)

XX DE Human tankyrase2 clone consensus protein SEQ ID NO: 101.

XX KW Human: tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
inflammatory disorder.

XX OS Homo sapiens.

XX PN WO200100849-A1.

XX PD 04-JAN-2001.

XX PF 28-JUN-2000; 2000WO-US17827.

XX PR 29-JUN-1999; 99US-0141582.

XX PA (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPT: 2001-102896/11.

XX N-PSDB; AAF63926.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by
poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
inflammatory and autoimmune disorders -

XX Example 1; Page 162-1665; 242pp; English.

XX The present invention provides the protein and coding sequence for the
human tankyrase2 protein. This is found in two different versions,
designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
polyADP-ribosylation activity and is involved in the modification of
TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
cancer. The sequences are useful in the treatment of cancers and
inflammatory disorders.

XX SQ Sequence 1169 AA;

Query Match 94.7%; Score 6121; DB 22; Length 1169;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 ARMSGRRRCAGGGAACASAAAAEVEPAARELFACRNGDVERVRLVTPKVNSRDTAGR 131
DB 1 ARMSGRRRCAGGGAACASAAAAEVEPAARELFACRNGDVERVRLVTPKVNSRDTAGR 60

QY 132 KSTPLHFAAGFRKDVVEYLLQNGANVQARDGGLIPLHNACSFHAEVNVNLLLRHGADP 191
DB 61 KSTPLHFAAGFRKDVVEYLLQNGANVQARDGGLIPLHNACSFHAEVNVNLLLRHGADP 120

QY 192 NARDNNWYTPHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYK 251
DB 121 NARDNNWYTPHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYK 180

QY 252 KDELLESARSNEEKMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADV 311
DB 181 KDELLESARSNEEKMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADV 240

QY 312 HAKDKGDLVPLHNACSYGHEVTELLVKGACVNMDLWQFTPLHEAASKNRVEVCSLLL 371

```
|||||
Db 241 HAKDKGLVPLHNACSYGHYEYTELLVKGACVNMADLWQFTPLHEAASKNVEVCSLLL 300
QY 372 SYGADPTLLNCHNKSAIDLAFTPQPKERLAYEFKGHSLLQAAREADVTRIKKHSLEWVN 431
Db 301 SYGADPTLLNCHNKSAIDLAFTPQPKERLAYEFKGHSLLQAAREADVTRIKKHSLEWVN 360
QY 432 FKHPOTHETALHCAASAPYPRKQICELLRLKGANINEKTEFLPLHVASEKAHNDVVE 491
Db 361 FKHPOTHETALHCAASAPYPRKQICELLRLKGANINEKTEFLPLHVASEKAHNDVVE 420
QY 492 VVVKHEAKYNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLQGFTALQMGNEV 551
Db 421 VVVKHEAKYNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLQGFTALQMGNEV 480
QY 552 QQLLEGISLGNSEADROLLEAAKAGDVETVKKLTQVSQVNCRDIEGROSTPLHFAAGYN 611
Db 481 QQLLEGISLGNSEADROLLEAAKAGDVETVKKLTQVSQVNCRDIEGROSTPLHFAAGYN 540
QY 612 RYSVVEYLLQHCADVHAKDKGLVPLHNACSYGHYEYAEVLVKGAVNVNADLWKFTPLH 671
Db 541 RYSVVEYLLQHCADVHAKDKGLVPLHNACSYGHYEYAEVLVKGAVNVNADLWKFTPLH 600
QY 672 EAAAKGVEICKLQLLQHGADPTKKNRDGNTPLDLVKDGDITDQLLGRDAALLDAAKKGC 731
Db 601 EAAAKGVEICKLQLLQHGADPTKKNRDGNTPLDLVKDGDITDQLLGRDAALLDAAKKGC 660
QY 732 LARVKKLSPDNVNCRDITQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAODKGLIPLH 791
Db 661 LARVKKLSPDNVNCRDITQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAODKGLIPLH 720
QY 792 NAAASYGHVDVAALLIKYNACVNATDKWFTPLHEAAKQRTQICALLAHGADPTLKNOE 851
Db 721 NAAASYGHVDVAALLIKYNACVNATDKWFTPLHEAAKQRTQICALLAHGADPTLKNOE 780
QY 852 GOTPLDLVSADVSALLTAAMPSPALPSYCKEQVLNGVRSFGATADALSSGSPSSSLSA 911
Db 781 GOTPLDLVSADVSALLTAAMPSPALPSYCKEQVLNGVRSFGATADALSSGSPSSSLSA 840
QY 912 ASSLDNLGSFSELSVSSSGTEGASSLEKEVPGVDFSIQFVRNLGLEHLMDFERE 971
Db 841 ASSLDNLGSFSELSVSSSGTEGASSLEKEVPGVDFSIQFVRNLGLEHLMDFERE 900
QY 972 QITLDVLVEMGHKEIKGINAYGHRHKLKIGVERLISQOQGLNPYLTLNTSGSCTILID 1031
Db 901 QITLDVLVEMGHKEIKGINAYGHRHKLKIGVERLISQOQGLNPYLTLNTSGSCTILID 960
QY 1032 LSPDDKEFOSVEEEMQSTVREHRDGGHAGGIENRNYNLKIQVCNKKLWERYTHRRKEYS 1091
Db 961 LSPDDKEFOSVEEEMQSTVREHRDGGHAGGIENRNYNLKIQVCNKKLWERYTHRRKEYS 1020
QY 1092 EENHNHANERMLFHGSPFVNALIIHKGFDERHAYIGMFCAGIYFAENSCKSNQYVYGIG 1151
Db 1021 EENHNHANERMLFHGSPFVNALIIHKGFDERHAYIGMFCAGIYFAENSCKSNQYVYGIG 1080
QY 1152 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALA 1211
Db 1081 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALA 1140
QY 1212 EYVYRGEQAYPEYLITYQIMRPEGMVDG 1240
Db 1141 EYVYRGEQAYPEYLITYQIMRPEGMVDG 1169
```

RESULT 6

AAB66295

ID AAB66295 standard; Protein; 1166 AA.

XX

AC AAB66295;

XX

DT 05-APR-2001 (first entry)

XX

DE Human tankyrase2 TANK2-SHORT SEQ ID NO: 135.

```
XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
XX Homo sapiens.
OS
PN WO200100849-A1.
XX
XX 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17827.
PR
XX 29-JUN-1999; 99US-0141582.
XX
PA (ICOS-) ICOS CORP.
XX
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
PI WPI; 2001-102896/11.
XX
XX N-PSDB; AAF63953.
DR
XX
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX
XX Claim 3; Page 200-203; 242pp; English.
XX
XX The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to aging and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
XX Sequence 1166 AA;
```

Query Match 94.5%; Score 6108; DB 22; Length 1166;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 75 MSGRRCAGGACACASAAAEVPAARELFEACRNGDVERVKRLVTEPKVNSRDTAGRKST 134
Db 1 MSGRRCAGGACACASAAAEVPAARELFEACRNGDVERVKRLVTEPKVNSRDTAGRKST 60
QY 135 PLHFAAGFRKDVVEYLLQNGANVOARDGGLPLHNACSFHAEVYVNLRLRHGADPNAR 194
Db 61 PLHFAAGFRKDVVEYLLQNGANVOARDGGLPLHNACSFHAEVYVNLRLRHGADPNAR 120
QY 195 DNWNYTPLHEAATKGKIDYCVILLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKKDE 254
Db 121 DNWNYTPLHEAATKGKIDYCVILLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKKDE 180
QY 255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK 314
Db 181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK 240
QY 315 DKGDVLPLHNACSYGHYEYTELLVKGACVNMADLWQFTPLHEAASKNVEVCSLLLSYG 374
Db 241 DKGDVLPLHNACSYGHYEYTELLVKGACVNMADLWQFTPLHEAASKNVEVCSLLLSYG 300
QY 375 ADPTLLNCHNKSAIDLAFTPQPKERLAYEFKGHSLLQAAREADVTRIKKHSLEWVNFKH 434
Db 301 ADPTLLNCHNKSAIDLAFTPQPKERLAYEFKGHSLLQAAREADVTRIKKHSLEWVNFKH 360
QY 435 PQTHETALHCAASAPYPRKQICELLRLKGANINEKTEFLPLHVASEKAHNDVVEVYV 494
Db 361 PQTHETALHCAASAPYPRKQICELLRLKGANINEKTEFLPLHVASEKAHNDVVEVYV 420
QY 495 KHEAKYNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLQGFTALQMGNEVQOL 554
Db 421 KHEAKYNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLQGFTALQMGNEVQOL 480
```

[illegible]

181	DB	LL	EAKSGN	SEKMMALITPL	LVNCHASDGRKSTPUHLAAGTNRVAIVULLLQHQRGADVAK	240
315	QY	DG	DLVPLH	NACSYGCHYEVT	TELLVKGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG	374
241	DB	DG	DLVPLH	NACSYGCHYEVT	TELLVKGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG	300
375	QY	AD	PTLLNCH	NKSAIDLAPT	POLKERLAYEFGHSLQLQAAREADVTRIKKHLSLEWNEFKH	434
301	DB	AD	PTLLNCH	NKSAIDLAPT	POLKERLAYEFGHSLQLQAAREADVTRIKKHLSLEWNEFKH	360
435	QY	PQ	THETALH	CAASPYPRKQ	ICEILLKRGANINEKTFEFTPLHVASEKAHNDVEVVV	494
361	DB	PQ	THETALH	CAASPYPRKQ	ICEILLKRGANINEKTFEFTPLHVASEKAHNDVEVVV	420
495	QY	KHEAKVN	ALDNLG	QTSLHRAAYC	GHLOTCRLLLSYGCDPNIISLOGFTALQMGNEVQOL	554
421	DB	KHEAKVN	ALDNLG	QTSLHRAAYC	GHLOTCRLLLSYGCDPNIISLOGFTALQMGNEVQOL	480
555	QY	LOEGIS	LGNS	EADQOLLEAAK	AGDVETVKLCTVQSVNCRDIEGRQSTPLHFAAGYNRV	614
481	DB	LOEGIS	LGNS	EADQOLLEAAK	AGDVETVKLCTVQSVNCRDIEGRQSTPLHFAAGYNRV	540
615	QY	VVEVIL	LOHCA	DVHAKD	KGGLVPLHNACSYGHEVAELLVKHGAVVYNADLWKFTPLHEAA	674

Db 541 VVEYLLQHGADVHAKDGGGLVPLHNAACSYGHYEVAELLVKHGAUVNVADLWKFTPLHEAA 600
QY 675 AKGYEICKLLQHGADPTKKNRDNTPDLVDKGDITDIDLLRGDAALLDAAKKGLAR 734
Db 601 AKGYEICKLLQHGADPTKKNRDNTPDLVDKGDITDIDLLRGDAALLDAAKKGLAR 660
QY 735 VKLSSPDVNCNCRDQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNA 794
Db 661 VKLSSPDVNCNCRDQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNA 720
QY 795 SYGHVDVAALLIKYNACVNATDKWATPLHEAAQKQRTQCALLLAHGADPTLKNQEGQT 854
Db 721 SYGHVDVAALLIKYNACVNATDKWATPLHEAAQKQRTQCALLLAHGADPTLKNQEGQT 780
QY 855 PLDLVADDDYVALLTAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 914
Db 781 PLDLVADDDYVALLTAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 840
QY 915 LDNLGSFSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVRLNGLHLMDFEREQIT 974
Db 841 LDNLGSFSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVRLNGLHLMDFEREQIT 900
QY 975 LDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISQOGLNPLYTLTNTSGSGTILDLSP 1034
Db 901 LDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISQOGLNPLYTLTNTSGSGTILDLSP 960
QY 1035 DKEFQSVSEEMOSTVREHRDGHAGGIFNRYNLIKQVCNKKLWERYTHRRKEYSEEN 1094
Db 961 DKEFQSVSEEMOSTVREHRDGHAGGIFNRYNLIKQVCNKKLWERYTHRRKEYSEEN 1020
QY 1095 HNANERMLFHGSPFFVNAIHKGFDERHAYIGMFGAGIYFAENSCKSNQYVYGIGGGTG 1154
Db 1021 HNANERMLFHGSPFFVNAIHKGFDERHAYIGMFGAGIYFAENSCKSNQYVYGIGGGTG 1080
QY 1155 CPVHKRSCYICHRQLLFCRVTLGKSLQFSAKMAHSPGHHSVTGRPSVNLALAEYV 1214
Db 1081 CPVHKRSCYICHRQLLFCRVTLGKSLQFSAKMAHSPGHHSVTGRPSVNLALAEYV 1140
QY 1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
Db 1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166

RESULT 8
AAY72589
ID AAY72589 standard; Protein; 1166 AA.
AC AAY72589;
XX AAY72589;

DT 02-MAY-2001 (first entry)

XX Human tankyrase homolog protein (THP).

DE Human; tankyrase homolog protein; THP: gene therapy; cancer;
KW tumour; basal cell carcinoma; therapy; genetic mapping;
KW cytostatic.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Misc-difference 368 /label= Met, Leu, Val
FT Misc-difference 392 /label= Asn, Thr
FT Misc-difference 415 /label= Val, Ile
FT Misc-difference 558 /label= Gly, Glu
FT Misc-difference 559 /label= Gly, Ala
FT Misc-difference 764 /label= Leu, Val
FT Misc-difference 884

FT XX /label= Asn, His, Asp, Tyr
PN WO200104326-A1.
XX 18-JAN-2001.
XX 03-JUL-2000; 2000WO-EP06609.
XX 09-JUL-1999; 99US-0350982.
XX (PHAA) PHARMACIA & UPJOHN SPA.
XX Berthelsen J, Toma S, Isacchi A;
XX WPI: 2001-168422/17.
XX N-PSDB; AAD02578.
XX New tankyrase homolog protein (THP) polynucleotide and polypeptide
XX useful in gene therapy, diagnosis and treatment or prevention of
XX unregulated cell growth, such as cancer or tumor cell growth -
XX Claim 20; Page 50-54; 60pp; English.
XX The present sequence is human tankyrase homolog
XX protein (THP). The THP polypeptides and polynucleotides of the
XX invention are useful in gene therapy and for treating or preventing
XX unregulated cell growth such as cancer or tumour (e.g. basal cell
XX carcinoma). The nucleic acid molecules of the invention and their
XX fragments are useful for restriction fragment length polymorphism
XX (RFLP) associated with certain disorders, as well as for genetic
XX mapping. Antisense oligonucleotides, or fragments of nucleic acid
XX encoding THP are useful as diagnostic tools for probing the
XX expression of Thp gene in various tissues. THP can be used as
XX antigens for raising antibodies against them and in assays for
XX identifying compounds that modulate their activity. They are used in
XX the manufacture of a medicament directed towards cancers or tumours.
XX THP are also useful for screening compounds in a variety of drugs
XX screening techniques and as a research tool for identification,
XX characterisation and purification of interacting, regulatory proteins.
XX SQ Sequence 1166 AA;

Query Match 93.6%; Score 6053; DB 22; Length 1166;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1157; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 75 MSRRRCAGGAGCAACASAAAEPAARELFEACRNGDVERVKRLVTEPKVNSRDTAGRKST 134
Db 1 MSRRRCAGGAGCAACASAAAEPAARELFEACRNGDVERVKRLVTEPKVNSRDTAGRKST 60
QY 135 PLHFAAGFGKRDVVEYLLONGANVOARDGGLIPLHNAACSYGHYEVAELLVKHGAUVNVADLWKFTPLHEAA 194
Db 61 PLHFAAGFGKRDVVEYLLONGANVOARDGGLIPLHNAACSYGHYEVAELLVKHGAUVNVADLWKFTPLHEAA 120
QY 195 DNWNVTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKKDE 254
Db 121 DNWNVTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKKDE 180
QY 255 LLESARSNGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNA 314
Db 181 LLESARSNGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNA 240
QY 315 DKGDLVPLHNAACSYGHYEYVELLVKHGACVNAAMDQWFTPLHEAAKSNRVEVCSLLLSYG 374
Db 241 DKGDLVPLHNAACSYGHYEYVELLVKHGACVNAAMDQWFTPLHEAAKSNRVEVCSLLLSYG 300
QY 375 ADPTLLNCHNKSAIDLAPTQPKERLAYEFKGHSLQAAREADVTRIKKHLSELMYNFKH 434
Db 301 ADPTLLNCHNKSAIDLAPTQPKERLAYEFKGHSLQAAREADVTRIKKHLSELMYNFKH 360
QY 435 POTHETALHCAAAASPYPKRKQICELLRLKGAXINEXTKKEFLPLHVASEKAHNDXVEVVV 494
Db 361 POTHETALHCAAAASPYPKRKQICELLRLKGAXINEXTKKEFLPLHVASEKAHNDXVEVVV 420


```
QY 495 KHEAKYNALDNLGOTSLHRAAYCGHLCQTRCLLSYGCDDPNIISLOGFTALQMGNEVQOL 554
DB 421 KHEAKYNALDNLGOTSLHRAAYCGHLCQTRCLLSYGCDDPNIISLOGFTALQMGNEVQOL 480
QY 555 LOEGISLGNSEADROLLLEAAKAGDVETVKLCITVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
DB 481 LOEGISLGNSEADROLLLEAAKAGDVETVKLCITVQSVNCRDIEGRQSTPLHFAAGYNRVS 540
QY 615 VVEYLQHGADYHAKDGLVPLHNACSYGHVEAEALLVKHGVNADLWKFTPLHFAA 674
DB 541 VVEYLQHGADYHAKDGLVPLHNACSYGHVEAEALLVKHGVNADLWKFTPLHFAA 600
QY 675 AKGKEICKLLQLHGADPTKKNRDNTPLDLVKDGDITDQLLRGDAALLDAAKKGCCLAR 734
DB 601 AKGKEICKLLQLHGADPTKKNRDNTPLDLVKDGDITDQLLRGDAALLDAAKKGCCLAR 660
QY 735 VKLSSPDVNVCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADYNAQDKGLIPLHNA 794
DB 661 VKLSSPDVNVCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADYNAQDKGLIPLHNA 720
QY 795 SVGHVDVAALLIKYNACVNATDKWAFTPLHFAAQKRGTLQCALILAHGADPTLKNQEGQT 854
DB 721 SVGHVDVAALLIKYNACVNATDKWAFTPLHFAAQKRGTLQCALILAHGADPTLKNQEGQT 780
QY 855 PLDLVSADDSVALLTAAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 914
DB 781 PLDLVSADDSVALLTAAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 840
QY 915 LNLSSFSFELSVMSSSTEGASLEKKEVPCVDFSTOFVRNIGLHLMDFIREQIT 974
DB 841 LNLSSFSFELSVMSSSTEGASLEKKEVPCVDFSTOFVRNIGLHLMDFIREQIT 900
QY 975 LDVLVPMGHKELKEIGINAYGHRHKLKIGVERLISGQQGLNPLTLNTSGSGTILIDLSP 1034
DB 901 LDVLVPMGHKELKEIGINAYGHRHKLKIGVERLISGQQGLNPLTLNTSGSGTILIDLSP 960
QY 1035 DKEFQSVSEEMQSVREHRDGGHAGGIFPNRYNLIKIOKVCNKKLWERYTHRRKVESEEN 1094
DB 961 DKEFQSVSEEMQSVREHRDGGHAGGIFPNRYNLIKIOKVCNKKLWERYTHRRKVESEEN 1020
QY 1095 HNHNANERMLFHGSPFVNALIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYIGGGTG 1154
DB 1021 HNHNANERMLFHGSPFVNALIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYIGGGTG 1080
QY 1155 CPVHKDRSYCHIRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYV 1214
DB 1081 CPVHKDRSYCHIRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYV 1140
QY 1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
DB 1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166

RESULT 9
AA97748
ID AA97748 standard; Protein: 1100 AA.
XX
AC AA97748;
XX
DT 06-AUG-2001 (first entry)
XX
DE Tankyrase homologue isotype 1 protein sequence.
XX
KW Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine.
XX
OS Unidentified.
XX
PN WO200130987-A2.
XX
PD 03-MAY-2001.
```

```
XX 25-OCT-2000; 2000WO-US41528.
XX
XX 25-OCT-1999; 99US-0427154.
XX
XX (RIGE-) RIGEL PHARM INC.
XX
XX Luo Y, Chan E, Xu X, Huang B;
XX
XX WPI: 2001-300503/31.
XX N-PSDB; AAA91487.
XX
XX Novel recombinant cell cycle polypeptide, tankyrase H useful for
XX inducing or preventing cell proliferation in cells, and for diagnosing,
XX treating or preventing cell cycle associated disorders such as cancer
XX
XX Claim 22; Fig 3: 63pp; English.
XX
XX This sequence is the Tankyrase homologue isotype 1 (TaHo-1) protein
XX of the invention. The invention also relates to the TaHo-2 protein.
XX The TaHo proteins are useful for inducing or preventing cell
XX proliferation in cells, and in the study or treatment of conditions
XX mediated by the cell cycle proteins, such as to diagnose, treat or
XX prevent cell cycle associated disorders, preferably cancer. The TaHo
XX coding sequences are useful as hybridisation probes, in chromosome and
XX gene mapping and in the generation of anti-sense DNA and RNA. The coding
XX sequences are also useful for the preparation of TaHo, for generating
XX either transgenic animals or knock out animals which, in turn, are useful
XX in a development and screening of therapeutically useful agents, in gene
XX therapy, as vaccine, and for construction of hybridisation probes for
XX mapping the gene which encodes TaHo and for the genetic analysis of
XX individuals with genetic disorders. The TaHo proteins, and their coding
XX sequences are useful in screening assays.
XX
XX Sequence 1100 AA:
XX
Query Match 89.2%; Score 5766; DB 22; Length 1100;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 141 GFGKRDVVEYLLQGANVQARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNYIT 200
DB 1 GFGKRDVVEYLLQGANVQARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNYIT 60
QY 201 PLHEAAIKGIDVCLVLLQHGAEPTIRNTDGTALDLPADPSAKAVLTGEYKDELLESAR 260
DB 61 PLHEAAIKGIDVCLVLLQHGAEPTIRNTDGTALDLPADPSAKAVLTGEYKDELLESAR 120
QY 261 SGNEEKMALLTPLNVNCHASDGRKRSTPLHLAAGYNRVKIVOLLQHGADYHAKKGDLY 320
DB 121 SGNEEKMALLTPLNVNCHASDGRKRSTPLHLAAGYNRVKIVOLLQHGADYHAKKGDLY 180
QY 321 PLHNACSYGHVEYVELLYKHGACYNAMDLPQFTPLHEAASKNRVEVCSLLSYGADPTLL 380
DB 181 PLHNACSYGHVEYVELLYKHGACYNAMDLPQFTPLHEAASKNRVEVCSLLSYGADPTLL 240
QY 381 NCHNKSADLAPTQPKERLAYEFKGHSLLOAAREADVTRIKKHLSEMVFVKHPQTHET 440
DB 241 NCHNKSADLAPTQPKERLAYEFKGHSLLOAAREADVTRIKKHLSEMVFVKHPQTHET 300
QY 441 ALHCAASAPYPRKQICELLRLKRGANINEKTEFTPLHVASAKAHNDVVEVVKHEAKV 500
DB 301 ALHCAASAPYPRKQICELLRLKRGANINEKTEFTPLHVASAKAHNDVVEVVKHEAKV 360
QY 501 NALDNLGOTSLHRAAYCGHLCQTRCLLSYGCDDPNIISLOGFTALQMGNEVQOLLEGIS 560
DB 361 NALDNLGOTSLHRAAYCGHLCQTRCLLSYGCDDPNIISLOGFTALQMGNEVQOLLEGIS 420
QY 561 LGNSEADROLLLEAAKAGDVETVKLCITVQSVNCRDIEGRQSTPLHFAAGYNRVSWEYLL 620
DB 421 LGNSEADROLLLEAAKAGDVETVKLCITVQSVNCRDIEGRQSTPLHFAAGYNRVSWEYLL 480
```

QY 621 QHGADVHAKDKGGGLVPLHNAACSYGHHYEVAAELLVYKHGAVVNVADLWKFTPLHEAAAKGKYE 680
 DB 481 QHGADVHAKDKGGGLVPLHNAACSYGHHYEVAAELLVYKHGAVVNVADLWKFTPLHEAAAKGKYE 540
 QY 681 ICKLLQHGADPTKKNRDNTPDLVKGDDTDIDLLRGDAALLDAAKKGCLARVKKLSS 740
 DB 541 ICKLLQHGADPTKKNRDNTPDLVKGDDTDIDLLRGDAALLDAAKKGCLARVKKLSS 600
 QY 741 PDVNCRDTOGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800
 DB 601 PDVNCRDTOGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
 QY 801 VAALLIKYNACVNATDKWATPLHEAAQKQRTQICALLLAHAGADPTLKNOGGOTPLDLVS 860
 DB 661 VAALLIKYNACVNATDKWATPLHEAAQKQRTQICALLLAHAGADPTLKNOGGOTPLDLVS 720
 QY 861 ADDVSALLTAAMPSPALPCYKQVNLGVRSFGATADALSSGSPSSLSAASLDNLG 920
 DB 721 ADDVSALLTAAMPSPALPCYKQVNLGVRSFGATADALSSGSPSSLSAASLDNLG 780
 QY 921 SFSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVFNGLGLEHLMDFEREQITLDVLVE 980
 DB 781 SFSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVFNGLGLEHLMDFEREQITLDVLVE 840
 QY 981 MGHKELKEIGINAYGHRHKLKIGVERLISQOGLNPYLTNTSGSGTILIDLPDDKEFQ 1040
 DB 841 MGHKELKEIGINAYGHRHKLKIGVERLISQOGLNPYLTNTSGSGTILIDLPDDKEFQ 900
 QY 1041 SVEEEMQSTVREHRDGGHAGGIFENRYNLIKIQVCNKKLWERTHRRKEVSEENHNHANE 1100
 DB 901 SVEEEMQSTVREHRDGGHAGGIFENRYNLIKIQVCNKKLWERTHRRKEVSEENHNHANE 960
 QY 1101 RMLFHSGSPFVNAIIHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYIGGGTGCPVHKD 1160
 DB 961 RMLFHSGSPFVNAIIHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYIGGGTGCPVHKD 1020
 QY 1161 RSCYICHRQLLCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYVYIRGEQ 1220
 DB 1021 RSCYICHRQLLCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYVYIRGEQ 1080
 QY 1221 AYPEYLITYQIMRPEGMDVG 1240
 DB 1081 AYPEYLITYQIMRPEGMDVG 1100

RESULT 10
 AAY05734

ID AAY05734 standard; Protein: 1074 AA.

AC AAY05734;
 XX

DT 19-JUL-1999 (first entry)

XX Human Grb7 effector 2.2412 protein.

DE Grb7 effector; 2.2412 protein; human; signal transduction;
 KW tumour marker; breast cancer; prostate cancer; prognosis;
 KW diagnosis.

XX Homo sapiens.

XX W09915647-A1.

XX 01-APR-1999.

XX 23-SEP-1998; 98WO-AU00795.

XX 23-SEP-1997; 97AU-0009388.

XX (GARV-) GARVAN INST MEDICAL RES.

PI Daly RJ, Sutherland RL;

XX

DR WPI; 1999-254707/21.
 DR N-PSDB; AAX25366.

XX New candidate effector for the Grb7 family of signalling proteins,
 PT and specific antibody, useful for detection and treatment of cancer
 PS Claim 9; Fig 1; 24pp; English.

XX The present sequence represents a novel candidate effector for
 CC the Grb7 family of signalling proteins, termed 2.2412. The
 CC sequence is predicted from a partial cDNA (see AAX25366). 2.2412
 CC mRNA transcripts (7 kb) were detected in all tissues examined
 CC with the exception of kidney. Expression was particularly high in
 CC skeletal muscle and placenta. Analysis of the sequence revealed
 CC significant homology to a large number of proteins containing
 CC ankyrin-like repeats. The 2.2412 gene was localised to between
 CC chromosome 10q33.2 and proximal 10q33.32. Deletions in the
 CC 10q22-25 region have been detected in human breast, prostate,
 CC renal, small cell lung and endometrial carcinomas, glioblastoma
 CC multiforme, melanoma and meningiomas. Detection of the protein
 CC encoded by the 2.2412 cDNA in a sample should provide a useful
 CC tumour marker and/or prognostic indicator for certain human
 CC cancers, in particular breast cancer and prostate cancer.
 CC Antagonism of the interaction between Grb7 family members and the
 CC encoded protein should provide a novel treatment strategy for human
 CC diseases exhibiting aberrant receptor tyrosine kinase signalling,
 CC such as cancer. Anti-Grb7 antibodies can be used in methods of
 CC detecting the presence of 2.2412 protein in a sample.

XX Sequence 1074 AA:

Query Match 86.7%; Score 5605; DB 20; Length 1074;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1069; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 167 IPLHNACSFGEAEVNNLLLRHGADPNARNWNNTPLHEAAIKGKIDVICIVLLOHGAEPTI 226
 DB 1 IPLHNACSFGEAEVNNLLLRHGADPNARNWNNTPLHEAAIKGKIDVICIVLLOHGAEPTI 60
 QY 227 RNTDGTALDADPSAKAVLTGEYKKDELESARSGNEFKMALLTPLNVNCHASDGRKS 286
 DB 61 RNTDGTALDADPSAKAVLTGEYKKDELESARSGNEFKMALLTPLNVNCHASDGRKS 120
 QY 287 TPLHLAAGYNRVKIVOLLQHGADVHAKDKGDLVPLHNACSYGHHYEVTELLVXHGACVNA 346
 DB 121 TPLHLAAGYNRVKIVOLLQHGADVHAKDKGDLVPLHNACSYGHHYEVTELLVXHGACVNA 180
 QY 347 MDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTQPKERLAYEFKG 406
 DB 181 MDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTQPKERLAYEFKG 240
 QY 407 HSLLOAAREADVTRIKKHLSELMVNFKHPOTHETALHCAASVPKPKQICELLLRKGAN 466
 DB 241 HSLLOAAREADVTRIKKHLSELMVNFKHPOTHETALHCAASVPKPKQICELLLRKGAN 300
 QY 467 INEKTKEFTPLHVASEKAHNDVVVVVHKEAKVNALDNLGQTSLHRAAYCGHLQTCRL 526
 DB 301 INEKTKEFTPLHVASEKAHNDVVVVVHKEAKVNALDNLGQTSLHRAAYCGHLQTCRL 360
 QY 527 LSYGCDPNIIISLOGFTALQNGENVQOLQEGISLGNSEADVQLLEAAKAGDVETVKKLC 586
 DB 361 LSYGCDPNIIISLOGFTALQNGENVQOLQEGISLGNSEADVQLLEAAKAGDVETVKKLC 420
 QY 587 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHHY 646
 DB 421 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHHY 480
 QY 647 EVAELLVKGAVNVADLWKFTPLHEAAAKGKYEICKLLQHGADPTKKNRDNTPLDLV 706
 DB 481 EVAELLVKGAVNVADLWKFTPLHEAAAKGKYEICKLLQHGADPTKKNRDNTPLDLV 540
 QY 707 KUGDTDIDQLLRGDAALLDAAKKGCLARVKKLSSPDVNCRDPTQGRHSTPLHLAAGYNNL 766
 DB 540 KUGDTDIDQLLRGDAALLDAAKKGCLARVKKLSSPDVNCRDPTQGRHSTPLHLAAGYNNL 766

Db 541 KGDGTDIQLRLLGDAALLDAKKGLARVKKLSSPDVNCVCRDTQGRHSTPLHLAAGYNNL 600
QY 767 EVAEYLLQHGADVNAODKGLPLHNAASYGHVDVAALLIKYNACVNAATDKWAFTPLHEA 826
Db 601 EVAEYLLQHGADVNAODKGLPLHNAASYGHVDVAALLIKYNASLNATDKWAFTPLHEA 660
QY 827 AOKGRTOLCALLAHGADPTLKNQEGOTPLDLVSADVDVSAALLTAAMPSPALPCYPOVL 886
Db 661 AOKGRTOLCALLAHGADPTLKNQEGOTPLDLVSADVDVSAALLTAAMPSPALPCYPOVL 720
QY 887 NGVSPGATADALSSGSPSSLSAASSLDNLGSGFSELSSVSSVSSGTEGASSLERKEVP 946
Db 721 NGVSPGATADALSSGSPSSLSAASSLDNLGSGFSELSSVSSVSSGTEGASSLERKEVP 780
QY 947 GVDFTSQFVRNLGLEHLMDFIREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVER 1006
Db 781 GVDFTSQFVRNLGLEHLMDFIREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVER 840
QY 1007 LISSQOGLNPYLTNTSGSGTILIDLSPPDKFQSVSEEMQSTVREHROGGHAGGIFNRY 1066
Db 841 LISSQOGLNPYLTNTSGSGTILIDLSPPDKFQSVSEEMQSTVREHROGGHAGGIFNRY 900
QY 1067 NILKIQVCNKKLWERYTHRRKEVSEENHNHNERMLFHGSPFVNALIIHKGFDERHAYIG 1126
Db 901 NILKIQVCNKKLWERYTHRRKEVSEENHNHNERMLFHGSPFVNALIIHKGFDERHAYIG 960
QY 1127 GMFGAGIYFAENSSKNQYVYGGTGPCVHKDRSICYIHRQLLFCRVTLGKSFQFSA 1186
Db 961 GMFGAGIYFAENSSKNQYVYGGTGPCVHKDRSICYIHRQLLFCRVTLGKSFQFSA 1020
QY 1187 MKMAHSPGPHSHSVTGRPSVNGALAEVYIRGEQAYPEYLITQIMRPEGWVDG 1240
Db 1021 MKMAHSPGPHSHSVTGRPSVNGALAEVYIRGEQAYPEYLITQIMRPEGWVDG 1074

RESULT 11

AAB47022

ID AAB47022 standard; Protein; 1166 AA.

XX AC AAB47022;

XX DT 29-MAR-2001 (first entry)

XX DE Human SPANK.

XX KW SPANK; SAM; sterile alpha motif; PARP; insulin resistance;

XX KW ANK; ankyrin repeat; cytosol; insulin-responsive aminopeptidase;

XX KW IRAP; GLUT4; adipocyte; insulin signalling pathway; hyperlipidaemia;

XX KW glucose intolerance; atheromatous disease; atherosclerosis;

XX KW obesity; cardiac insufficiency; coronary insufficiency; stroke;

XX KW high blood pressure; non-insulin dependent diabetes; hypertension;

XX KW hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy.

XX OS Homo sapiens.

XX FH Key

FT Misc-difference 173..209

FT /note= "Encoded by nucleotides 768..878"

FT Misc-difference 327..362

FT /note= "Encoded by nucleotides 1230..1337"

FT Misc-difference 375..398

FT /note= "Encoded by nucleotides 1374..1445"

FT Misc-difference 482..524

FT /note= "Encoded by nucleotides 1695..1823"

FT Misc-difference 641..677

FT /note= "Encoded by nucleotides 2172..2282"

XX WO200077225-A1.

XX 21-DEC-2000.

XX 09-JUN-2000; 2000WO-US15926.

XX 11-JUN-1999; 99US-0138957.
(WHED) WHITEHEAD INST BIOMEDICAL RES.
(GEHO) GEN HOSPITAL CORP.
XX Chi N, Lodish HF;
XX WPI; 2001-091404/10.
DR N-PSDB; AAC85294.
XX New insulin signalling protein SPANK, useful for reducing body mass,
XX glucose intolerance or insulin resistance and for preventing or
XX treating obesity-related and muscle-related diseases -
XX Claim 3; Fig 3; 65pp; English.
XX This sequence represents human SPANK. The SPANK protein comprises
XX 3 domains:
XX (a) a SAM (sterile alpha motif) domain;
XX (b) a PARP (poly adenosine diphosphate-ribose polymerase) catalytic
XX domain; and
XX (c) an ANK domain composed of ankyrin repeats
XX SPANK is a cytosolic protein which can poly(ADP-ribose)ylate itself.
XX SPANK binds insulin-responsive aminopeptidase (IRAP) and modulates
XX translocation of GLUT4 in the perinuclear region of adipocytes. It
XX is an effector in the insulin signalling pathway in eukaryotic cells.
XX SPANK is useful for reducing body mass, reducing glucose
XX intolerance or insulin resistance, for preventing or treating
XX obesity-related diseases or disorders, such as obesity, cardiac
XX insufficiency, coronary insufficiency, stroke, hypertension,
XX atheromatous disease, atherosclerosis, high blood pressure, non-insulin
XX dependent diabetes, hyperlipidaemia, hyperuricemia and Syndrome X and is
XX also useful for preventing or treating muscle-related diseases or
XX disorders, such as muscular dystrophy, muscle atrophy and muscle
XX fatigue. Antibodies immunospecific for SPANK are useful for detecting
XX the presence of SPANK polypeptide in a biological sample.
XX Sequence 1166 AA;

Query Match 85.8%; Score 5543; DB 22; Length 1166;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 1092; Conservative 0; Mismatches 0; Indels 148; Gaps 10;

QY 75 MSGRRCAGGAGCAASAAAEVPAARELFACRNGDVERVKRLVTPKVNRSRDAGRKST 134
Db 1 MSGRRCAGGAGCAASAAAEVPAARELFACRNGDVERVKRLVTPKVNRSRDAGRKST 60
QY 135 PLHFAAGFGRKDVVEYLLQNGANVQARDGGLIPLHNACSFHGAHVNNLLRHGADPNAR 194
Db 61 PLHFAAGFGRKDVVEYLLQNGANVQARDGGLIPLHNACSFHGAHVNNLLRHGADPNAR 120
QY 195 DNNYTPLHEAAIKGKTDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVL----- 246
Db 121 DNNYTPLHEAAIKGKTDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLKMALLTP 180
QY 247 -----TGEYKKDELLESARSNGNEKMMALLTPLNVNCHASDGRKSTPLHLAAGYN 296
Db 181 LNVNCHASDGTGEYKKDELLESARSNGNE-----RKSTPLHLAAGYN 222
QY 297 RVKIVQLLQHGADVHAKDKGLVPLHNACSYGHVEYVELLVKHGACVNAWDLQWFTPLH 356
Db 223 RVKIVQLLQHGADVHAKDKGLVPLHNACSYGHVEYVELLVKHGACVNAWDLQWFTPLH 282
QY 357 EAASKNRVEVCSSLLLSYGADPTLLNCHNKAIDLAPFPOLKERL----- 400
Db 283 EAASKNRVEVCSSLLLSYGADPTLLNCHNKAIDLAPFPOLKERLIKHLKLSLEWVNFHQP 342
QY 401 AYEFGKSHLQAAAREADVTRIKKHLKLSLEWVNFHQPHTALHCAAAAPYKPKQICELL 460
Db 343 AYEFGKSHLQAAAREADVTR-----THETALHCAAAAS---KRQICELL 383
QY 461 LRKGANINEKTK---EFLTPLHVASEKAHNDVVEVYVKEAKVNALDNLGQTSLHRAAYC 517

```
|||||
Db 384 LRKGANINEKTPYPPEFTPLHASEKANDVVEVVVHKAQNALDNLGQTSLHRAAYC 443
QY 518 GHLQTCRLLSYGCDPNITISLOGFTALOMGNENVOQLL-----QEGI 559
Db 444 GHLQTCRLLSYGCDPNITISLOGFTALOMGNENVOQLLTVKKLCIVQSVNCRDIEGQEI 503
QY 560 SLGNSEADRLLEAAKAGDVEVTKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYL 619
Db 504 SLGNSEADRLLEAAKAGDVE-----RQSTPLHFAAGYNRVSVVEYL 545
QY 620 LQHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKGAVNVVADLWKFTPLHEAAKAGKY 679
Db 546 LQHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKGAVNVVADLWKFTPLHEAAKAGKY 605
QY 680 EICKLLQHGADPTKKNRDNPTPLDVKDGDIDQDLLRGAALLDAKKGCLARVKKLS 739
Db 606 EICKLLQHGADPTKKNRDNPTPLDVKDGDIDQ-----RVKKLS 646
QY 740 SPDVNCRTDQ-----RHSTPLHLAAGYNLEVAEYLLQHGADVN 780
Db 647 SPDVNCRTDQDLLRGAALLDAKKGCLARHSTPLHLAAGYNLEVAEYLLQHGADVN 706
QY 781 AODKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFTPLHEAAKGRTOQLCALLA 840
Db 707 AODKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFTPLHEAAKGRTOQLCALLA 766
QY 841 HGADPTLKNQEGOTPLDVSADVDVSAALLTAAMPSPALPCYKPOVLNGVRSPGATADALS 900
Db 767 HGADPTLKNQEGOTPLDVSADVDVSAALLTAAMPSPALPCYKPOVLNGVRSPGATADALS 826
QY 901 SGPSPSSLSAASSLDNLGSGFSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVRNLG 960
Db 827 SGPSPSSLSAASSLDNLGSGFSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVRNLG 886
QY 961 LEHLMDFRBOITLDVLVEMGHKELKETGINAYGHRHKLKIGVERLISGQOGLNPYLTL 1020
Db 887 LEHLMDFRBOITLDVLVEMGHKELKETGINAYGHRHKLKIGVERLISGQOGLNPYLTL 946
QY 1021 NTSGSTILDLSPDKEFQSVSEEMOSTVREHRRDGGHAGGIFNRYNLIKIOKVCNKKLW 1080
Db 947 NTSGSTILDLSPDKEFQSVSEEMOSTVREHRRDGGHAGGIFNRYNLIKIOKVCNKKLW 1006
QY 1081 ERYTHRRKEVSEENHNHANERMLFHGSPFVNALIHKGFDERHAYIGGMFGAGIYFAENSS 1140
Db 1007 ERYTHRRKEVSEENHNHANERMLFHGSPFVNALIHKGFDERHAYIGGMFGAGIYFAENSS 1066
QY 1141 KSNQYVYIGGGTGPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHHSVT 1200
Db 1067 KSNQYVYIGGGTGPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHHSVT 1126
QY 1201 GRPSVNGLALAEYVYRGOAYPEYLITYQIMRPEGWVDG 1240
Db 1127 GRPSVNGLALAEYVYRGOAYPEYLITYQIMRPEGWVDG 1166
```

RESULT 12

ID AAB27210
XX AAB27210 standard; Protein; 1431 AA.

AC AAB27210;

DT 27-FEB-2001 (first entry)

XX Human tankyrase II protein sequence SEQ ID NO: 4.

DE Human; tankyrase II; telomere length; signal transduction.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 41 /label= Xaa

```
FT Misc-difference 97 /note= "encoded by ANG"  
FT /label= Xaa  
FT /note= "encoded by ANG"  
FT Misc-difference 109 /label= Xaa  
FT /note= "encoded by NCG"  
FT Misc-difference 113 /label= Xaa  
FT /note= "encoded by ANG"  
FT /label= Xaa  
FT /note= "ANK domain"  
FT Domain 122..935  
FT Misc-difference 125 /label= Xaa  
FT /note= "encoded by AAN"  
FT Misc-difference 138 /label= Xaa  
FT /note= "encoded by NTC"  
FT Misc-difference 165 /label= Xaa  
FT /note= "encoded by NGT"  
FT Misc-difference 172..173 /label= XaaXaa  
FT /note= "encoded by NTTRAA"  
FT Misc-difference 177 /label= Xaa  
FT /note= "encoded by CAN"  
FT Misc-difference 184 /label= Xaa  
FT /note= "encoded by NCA"  
FT Misc-difference 187 /label= Xaa  
FT /note= "encoded by NAT"  
FT Misc-difference 201 /label= Xaa  
FT /note= "encoded by CMT"  
FT Misc-difference 203..204 /label= XaaXaa  
FT /note= "encoded by RAAANT"  
FT Misc-difference 206 /label= Xaa  
FT /note= "encoded by NAT"  
FT Misc-difference 210 /label= Xaa  
FT /note= "encoded by NAM"  
FT Misc-difference 212 /label= Xaa  
FT /note= "encoded by CNA"  
FT /label= Xaa  
FT /note= "encoded by NGT"  
FT Misc-difference 214 /label= Xaa  
FT /note= "encoded by NAM"  
FT Misc-difference 226..227 /label= XaaXaa  
FT /note= "encoded by CNCNAT"  
FT Misc-difference 236..237 /label= XaaXaa  
FT /note= "encoded by GANNNT"  
FT Misc-difference 666 /label= Xaa  
FT /note= "encoded by TNT"  
FT Misc-difference 724 /label= Xaa  
FT /note= "encoded by AAN"  
FT Misc-difference 726 /label= Xaa  
FT /note= "encoded by CNT"  
FT Misc-difference 734 /label= Xaa  
FT /note= "encoded by GAN"  
FT Misc-difference 740 /label= Xaa  
FT /note= "encoded by NTC"  
FT Misc-difference 762..763
```

FT /label= XaaXaa
 FT /note= "encoded by TTNTNT"
 FT 947..993
 FT Domain
 FT /note= "SAM domain"
 FT Misc-difference 992
 FT /label= Xaa
 FT /note= "encoded by NAG"
 FT Misc-difference 1267
 FT /label= Xaa
 FT /note= "encoded by TAA"
 FT Misc-difference 1274
 FT /label= Xaa
 FT /note= "encoded by TAA"
 FT Misc-difference 1277
 FT /label= Xaa
 FT /note= "encoded by TGA"
 FT Misc-difference 1279
 FT /label= Xaa
 FT /note= "encoded by TAA"
 FT Misc-difference 1294
 FT /label= Xaa
 FT /note= "encoded by TGA"
 FT Misc-difference 1314
 FT /label= Xaa
 FT /note= "encoded by TAA"
 FT Misc-difference 1322..1323
 FT /label= XaaXaa
 FT /note= "encoded by TAATAA"
 FT Misc-difference 1353
 FT /label= Xaa
 FT /note= "encoded by TAA"
 FT Misc-difference 1358
 FT /label= Xaa
 FT /note= "encoded by TGA"
 FT Misc-difference 1371
 FT /label= Xaa
 FT /note= "encoded by TAA"
 FT Misc-difference 1377
 FT /label= Xaa
 FT /note= "encoded by TAA"
 FT Misc-difference 1380
 FT /label= Xaa
 FT /note= "encoded by TAA"
 FT Misc-difference 1393
 FT /label= Xaa
 FT /note= "encoded by TCA"
 FT Misc-difference 1399
 FT /label= Xaa
 FT /note= "encoded by TAA"
 FT Misc-difference 1403
 FT /label= Xaa
 FT /note= "encoded by TAA"
 FT Misc-difference 1408
 FT /label= Xaa
 FT /note= "encoded by TAG"
 FT Misc-difference 1419
 FT /label= Xaa
 FT /note= "encoded by TAA"
 FT /note= "Xaa-unknown"
 XX
 XX WO200061813-A1.
 XX
 XX 19-OCT-2000.
 PD
 XX 10-APR-2000; 2000WO-US09558.
 PF
 XX 09-APR-1999; 9905-0128577.
 PR
 PR 13-APR-1999; 9905-0129123.
 XX
 XX (GERO-) GERON CORP.
 PA
 XX Morin GB, Funk WD, Piatyszek MA;
 PI
 XX

DR WPI; 2000-679503/66.
 XX N-PSDB; AAC66824.
 XX
 PT Novel mammalian Tankyrase II polypeptide and the polynucleotide
 PT encoding the polypeptide useful for modulating or maintaining telomere
 PT length, replicative capacity, apoptosis, chromosome packing or gene
 PT expression -
 XX
 PS Disclosure; Fig 3; 52pp; English.
 XX
 CC The present sequence is a version of the human tankyrase II protein
 CC sequence. The protein is thought to be involved in signal transduction in
 CC the cell, and to have binding activity for other telomere-associated
 CC proteins. It is possible that it plays a role in the regulation of
 CC telomere length, thus affecting the replicative ability of the cell. The
 CC protein is useful for ribosylating target proteins, for determining
 CC tankyrase II binding activity in a sample, and for modulating telomere
 CC length in a cell.
 XX
 SQ Sequence 1431 AA;
 Query Match 84.7%; Score 5474; DB 21; Length 1431;
 Best Local Similarity 86.6%; Pred. No. 0;
 Matches 1076; Conservative 6; Mismatches 157; Indels 4; Gaps 2;
 QY 1 RCSARRGAAGGQAGQARGARVGAAGHGTADPDVPTAGSQAARALSASSPGGLALLAGPGLL 60
 DB 25 RCLRRGAAGGQAGHXRGCARGRGHGTADPDVPTAGSQAARALSASSPGGLALLAGPGLL 84
 QY 61 RLLALLLAVAAARIMSGRRCAAGGGGAAACASAAAEVPAARELFEACRNGDVERVKRLVTP 120
 DB 85 RLLALLLAVAAAXIMSGRRCAAGGGGACAXAAAEVPAARELFEACRNGDVERKKLVTP 144
 QY 121 EKVNSRDTAGRKSTPLHEAAGFGRKDVVEYLLONGANVQARDGGLIPLHNACSFGEAEV 180
 DB 145 EKVNSRDTAGRKSTPLHFPAXFGRKDLXXYLLXNGANVQXRDGGLIPLHNACSFGEAXX 204
 QY 181 VNLLLRHGADPNARDNNVYTPLEHAAIKGKIDVCIVLQHGAEPTIRNTDGTALDLADP 240
 DB 205 IXLLXHXAXPNARDNNVYTPXEEAAIKGKIXXCIVLQHGAEPTIRNTDGTALDLADP 264
 QY 241 SAKAVLTGEYKKDELLESARSNEEKMMALLTPLNVNCHASDGRKSTPLHAAAGVNRVKI 300
 DB 265 SAKAVLTGEYKKDELLESARSNEEKMMALLTPLNVNCHASDGRKSTPLHAAAGVNRVKI 324
 QY 301 VOLLLQHGADVHAKDKDGLVPLHNACSYCHYEVTETLLVKGACVNMADLWQFTPLHEAAS 360
 DB 325 VOLLLQHGADVHAKDKDGLVPLHNACSYCHYEVTETLLVKGACVNMADLWQFTPLHEAAS 384
 QY 361 KNRVEVCSLLLSYGADPTLLNCHNKSADLAPTOLKERLAYEFKGHSLLOAAREADVTR 420
 DB 385 KNRVEVCSLLLSYGADPTLLNCHNKSADLAPTOLKERLAYEFKGHSLLOAAREADVTR 444
 QY 421 IKKHLISLEMVNFKHPOTHETALHCAAAASPYPKRKOICEALLRKGANINEKTEFTPLHV 480
 DB 445 IKKHLISLEMVNFKHPOTHETALHCAAAASPYPKRKOICEALLRKGANINEKTEFTPLHV 504
 QY 481 ASEKAHNDVVEVVKHEAKVNALDNLGOTSLEHRAAYCGHLQTCRLLLSYGCDPNTIISLOG 540
 DB 505 ASEKAHNDVVEVVKHEAKVNALDNLGOTSLEHRAAYCGHLQTCRLLLSYGCDPNTIISLOG 564
 QY 541 FTALQMGNEVVOQLLOEGISLGNSEADROLLEAAKAGDVETVKKLTCTVQSVNCRDIEGRQ 600
 DB 565 FTALQMGNEVVOQLLOEGISLGNSEADROLLEAAKAGDVETVKKLTCTVQSVNCRDIEGRQ 624
 QY 601 STPLHFAAGYNRVSVVVEYLLQHGADVHAKDKGGLVPLHNACSYGHEYVAELLVKHGAAYN 660
 DB 625 STPLHFAAGYNRVSVVVEYLLQHGADVHAKDKGGLVPLHNACSYGHEYVAELLVKHGAAYN 684
 QY 661 VADLWKFTPLHEAAAKGKYEICKLLOHGADPTKKNRDNGTPLDLVKDGTDTIDQLLRGD 720
 DB 685 VADLWKFTPLHEAAAKGKYEICKLLOHGADPTKKNRDNGTPLDLVKDGTDTIDQLLRGD 744

QY 721 AALLDAARKGCLARYKLLSSPDVNCNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780
 Db 745 AVLLDAARKGCLARYKXXFXFPDNYNCRDTQGRHSTPLHL-AGXXXXXXXKXXXXXXX 803
 QY 781 AODKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFPLHEAAQKGTQLCAULLA 840
 Db 804 XXX 863
 QY 841 HGADPTLKNOBGOPLDLVSADDDYVALLTAAMPSPALPSCYKPOVLNGVRSPGATADALS 900
 Db 864 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXTAAAPPVLPSCNKQVNLNGVRSPGATADALS 923
 QY 901 SGPPSSLSAASLDNLGSGFSLSVSSGTEGASSLEKKEVPGVDFPSITOFVNLG 960
 Db 924 SGPPSSLSAASLDNLGSGFSLSVSSGTEGASSLEKKEVPGVDFPSITOFVNLG 983
 QY 961 LEHLMDFERQITDLVEMGHKELKEIGINAYGHRHKLKIGVERLISGOQGNPYLTL 1020
 Db 984 LEHLMDFXREQITDLVEMGHKELKEIGINAYGHRHKLKISFERLISGOQGNPYLTL 1043
 QY 1021 NTSGSGTILDLSPDDKEFQSVSEEMOSTVREHRDGGHAGGIFNRYNILKIQVCNK--- 1077
 Db 1044 NTSGSGTILDLSPDDKEFQSVSEEMOSTVREHRDGGHAGGIFNRYNILKIQVCNRAKI 1103
 QY 1078 KIWERYTHRRKEVSEENHNHANERMLERHSGSPFYNAIILHKGFDERHAYIGCMFGAGIYFAE 1137
 Db 1104 RHEBRYTHRRKEVSEENHNHANERMLERHSGSPFYNAIILHKGFDERHAYIGCMFGAGIYFAE 1163
 QY 1138 NSSKSNQVYVIGGGTGCPCVHKRSCYICHRQLLFCRVTLGKSPLOFSAMKMAHSPGCHH 1197
 Db 1164 NSSKSNQVYVIGGGTGCPCVHKRSCYICHRQLLFCRVTLGKSPLOFSAMKMAHSPGCHH 1223
 QY 1198 SVTGRPSVNGLAAEYVIYGEQAYPEYLITYQIMRPEGWVDG 1240
 Db 1224 SVTGRPSVNGLAAEYVIYGEQAYPEYLITYQIMRPEGWVDG 1266

RESULT 13

AAB27212

ID AAB27212 standard; Protein; 1327 AA.

XX AC AAB27212;

XX AC AAB27212;

DT 27-FEB-2001 (first entry)

XX Human tankyrase I protein sequence SEQ ID NO: 8.

XX Human; tankyrase II; telomere length; signal transduction.

KW Homo sapiens.

OS Wo200061813-A1.

XX PD 19-OCT-2000.

XX PF 10-APR-2000; 2000WO-US09558.

XX PR 09-APR-1999; 99US-0128577.

XX PR 13-APR-1999; 99US-0129123.

XX PA (GERO-) GERON CORP.

XX PI Morin GB, Funk WD, Piatyszek MA;

XX XX WPI; 2000-679503/66.

XX PT Novel mammalian Tankyrase II polypeptide and the polynucleotide

XX PT encoding the polypeptide useful for modulating or maintaining telomere

XX PT length, replicative capacity, apoptosis, chromosome packing or gene

XX PT expression

XX PS Claim 4; Fig 5; 52pp; English.

XX XX

CC The present sequence is a version of the human tankyrase I protein
 CC sequence. The invention relates to the isolation of the protein and
 CC coding sequences of human tankyrase II. This protein is thought to be
 CC involved in signal transduction in the cell, and to have binding activity
 CC for other telomere-associated proteins. It is possible that it plays a
 CC role in the regulation of telomere length, thus affecting the replicative
 CC ability of the cell. The protein is useful for ribosylating target
 CC proteins, for determining tankyrase II binding activity in a sample, and
 CC for modulating telomere length in a cell.

XX Sequence 1327 AA;

Query Match 79.0%; Score 5103.5; DB 21; Length 1327;
 Best Local Similarity 79.2%; Pred. No. 0;
 Matches 963; Conservative 110; Mismatches 128; Indels 15; Gaps 3;

QY 22 AAHGTADPVTAGSQAARALASASSPGLALLAGPGLLLRLLALLAVAAARIMSGRCA 81
 Db 112 SAAGVAPNPAGSGNNPSSSSPTSSSSSPSPG-----SSLAESPEAGVSYSTAPL 165

QY 82 GGAACASAAAAEVEPAARELFEACRNGDVERVKRLVTPPEKVNRSRDTAGRKSTPLHFAAG 141
 Db 166 GPCAAGPCTGVPAVSGALRELLEACRNGDYSVRKRLVDAANVNAKDMAGRKSSPLHFAAG 225

QY 142 FGRKDVVEYLLQNGANYOARDGGLIPLHNACSFHAEVNVNLLLRHGDADPNARDNNWYTP 201
 Db 226 FGRKDVVEYLLQNGANYOARDGGLIPLHNACSFHAEVNVNLLLRHGDADPNARDNNWYTP 285

QY 202 LHEAAIKGKIDVCLVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKDELLESARS 261
 Db 286 LHEAAIKGKIDVCLVLLQHGADPNIRNTDGTALDLADPSAKAVLTGEYKDELLESARS 345

QY 262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNNRVIVQLLQHGADVHAKDKGLVP 321
 Db 346 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNNRVIVQLLQHGADVHAKDKGLVP 405

QY 322 LHNACSYGHYEVTLLVKGACVNMALWQFTPLHEAASKNRVVCSSLISYGADPTLLN 381
 Db 406 LHNACSYGHYEVTLLVKGACVNMALWQFTPLHEAASKNRVVCSSLISYGADPTLLN 465

QY 382 CHNKSAITDAPTPOLKRLAYEFKSHLSLOAAREADYTRIKKHLSEWVNFKHPQTHETA 441
 Db 466 CHKSADMAPTPOLKRLAYEFKSHLSLOAAREADYTRIKKHLSEWVNFKHPQTHETA 525

QY 442 LHCAASPYPKRQICELLKRGANINEKTEFLTPPLHVAEKANNDVVEYVVKHEAKVN 501
 Db 526 LHCAASPYPKRQICELLKRGANINEKTEFLTPPLHVAEKANNDVVEYVVKHEAKVN 585

QY 502 ALDNLGOTSLHRAAYCCHLQTCRLLLSYGCDDPNIIISLQGTALQMGNEVQOQLQESL 561
 Db 586 ALDNLGOTSLHRAAYCCHLQTCRLLLSYGCDDPNIIISLQGTALQMGNEVQOQLQESL 645

QY 562 GNSEADROLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNNRVVVEYLLQ 621
 Db 646 RTSDVYRLLLEASKAGDLETVKOLCSSQYNNCRDIEGRQSTPLHFAAGYNNRVVVEYLLH 705

QY 622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAYVNVADLWKFPTPLHEAAAKGYEI 681
 Db 706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAYVNVADLWKFPTPLHEAAAKGYEI 765

QY 682 CKLLQHGADPTKKNRDGNTPLDLVKDGDPTDIDLLRGDAALLDAKKGCLARVKLSSP 741
 Db 766 CKLLQHGADPTKKNRDGNTPLDLVKDGDPTDIDLLRGDAALLDAKKGCLARVKLSSP 825

QY 742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801
 Db 826 ENINCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDI 885

QY 802 AALLIKYNACVNATDKWAFPLHEAAQKGTQLCAULLAHGADPTLKNQBGQPTPLDLVSA 861
 Db 886 AALLIKYNACVNATDKWAFPLHEAAQKGTQLCAULLAHGADPTLKNQBGQPTPLDLVSA 945

QY 862 DDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASLDNLGSGS 921

Db 886 AALLIKYNTCVNATDKWAFPLHEAAQGRQLCALLAHGADPTMKNQEQOTPLDLATA 945
 QY 862 DDVSAALLTAAMPSPALPSCYKQVNLNGVRSPGATADALSQSPSSLSAASSLDNLGSG 921
 Db 946 DDIRALLIDAMPPEALPTCFKPOAT-----VVSASLISPASTPSCLSAASSIDNLGTP 998
 QY 922 FSELSSVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDFIREQITLDLV 979
 Db 999 LAELAVGASNAGDGAACERKECEVAGLDWNTISQFLKSLGLEHLRDFITEQITLDVLA 1058
 QY 980 EMGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPLYLTNTSGSGTILIDSPDKDEF 1039
 Db 1059 DMGHEELKEIGINAYGHRHKLKIGVERLISGQOGLNPLYLTNTSGSGTILIDSPDKDEF 1118
 QY 1040 QSVVEEMQSTVREHRDGHAGGIFNRYNLIKQVCKNKLWERTHRKEVSEENHNHAN 1099
 Db 1119 QSVVEEMQSTVREHRDGHAGGIFNRYNLIKQVCKNKLWERTHRKEVSEENHNHAN 1178
 QY 1100 ERMFLHGSFVNALIIHKGDFDERHAYIGMFCAGIYFAENSCKSNQYVYGGTGCPCVHK 1159
 Db 1179 ERMFLHGSFVNALIIHKGDFDERHAYIGMFCAGIYFAENSCKSNQYVYGGTGCPCVHK 1238
 QY 1160 DRSCYICHRQLLCRVTLGKSFLOFSAMKWAHSPPGHSHSVTGRPSVNGIALAEVYIYRGE 1219
 Db 1239 DRSCYICHRQLLCRVTLGKSFLOFSAMKWAHSPPGHSHSVTGRPSVNGIALAEVYIYRGE 1298
 QY 1220 QAYPEYLITYQIMRPE 1235
 Db 1299 QAYPEYLITYQIMKPE 1314

RESULT 15

AAB66279
 ID AAB66279 standard; Protein; 1327 AA.

XX AAB66279;
 AC AAB66279;
 XX 05-APR-2001 (first entry)
 XX Human tankyrase1 SEQ ID NO: 4.
 DE Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasia; aging;
 KW Inflammatory disorder.
 KW Homo sapiens.
 OS WO200100849-A1.
 PN 04-JAN-2001.
 XX 28-JUN-2000; 2000WO-0517827.
 XX 29-JUN-1999; 99US-0141582.
 XX (ICOS-) ICOS CORP.
 PA Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
 PI WPI: 2001-102896/11.
 DR N-PSDB; AAF63838.
 XX
 PT New tankyrase2 polypeptides, useful for treating conditions mediated by
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
 PT inflammatory and autoimmune disorders -
 XX Example 1; Page 118-121; 242pp; English.
 PS
 XX The present invention provides the protein and coding sequence for the
 CC human tankyrase2 protein. This is found in two different versions,
 CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
 CC polyADP-ribosylation activity and is involved in the modification of
 CC TRF1, which is a telomere-specific binding protein. The regulation of

CC telomere length, in which TRF1 has a role, is linked to ageing and
 CC cancer. The sequences are useful in the treatment of cancers and
 CC inflammatory disorders.
 XX SQ Sequence 1327 AA;
 Query Match 79.0%; Score 5103.5; DB 22; Length 1327;
 Best Local Similarity 79.2%; Pred. No. 0;
 Matches 963; Conservative 110; Mismatches 128; Indels 15; Gaps 3;
 QY 22 AAHTAPDPVTVTAQAARALASASSPGGLALLAGPGLLRLLALLLAAVAARIMSGRRCA 81
 Db 112 SAAGVAPNPAGSGNNSSPSSSTSSSSSSPSPG-----SSLAESPEAAGVSTAPL 165
 QY 82 GGGACACAAAEAVEPAARELFEACRNGDVERVKRLVTPKEVNSRDTAGRKSTPLHFAAG 141
 Db 166 GPGAAGPGTGPVAVSGALRELEACRNGDVSVRKRLVDAANVNNAKDMAGKSSPLHFAAG 225
 QY 142 FGRKDVVEYLQNCANVOARDGGLIPLHNACSPGHAEVNVLLLRHGADPNARDNNWYTP 201
 Db 226 FGRKDVVEHLLQNCANVHARDGGLIPLHNACSPGHAEVNVLLLRHGADPNARDNNWYTP 285
 QY 202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLPDSAKAVLTGEYKDELLEARS 261
 Db 286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGSALDLPDSAKAVLTGEYKDELLEARS 345
 QY 262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNNRVIVOLLQHGADYHAKDGLVP 321
 Db 346 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNNRVIVOLLQHGADYHAKDGLVP 405
 QY 322 LHNACSYGHYEVELLYKVGACVNMADLWQFTPLHAAASKNRVEVCSLLLSYGADPTLLN 381
 Db 406 LHNACSYGHYEVELLYKVGACVNMADLWQFTPLHAAASKNRVEVCSLLLSYGADPTLLN 465
 QY 382 CHNKSALDAPTPQLKERLAYEPKSHSLQAARADYTRIKKHLSLEWFKHPQTHETA 441
 Db 466 CHGKSADMAPTPELRERLTPEFKHSLQAARADYTRIKKHLSLEWFKHPQTHETA 525
 QY 442 LHCAASPYKPKOICELLRLKGANINEKTEFTPLHVAASEKAHNDVNVVVKHEAKVN 501
 Db 526 LHCAVASLHPKPKOVTELLRLKGANVNEKNDFMTPPLHVAASEKAHNDVNVVVKHEAKVN 585
 QY 502 ALDNLGQTSLHRAAYCGHLOTCRLLLSYSGDPNIIISLQGTALQMGNNVQQLQEQEISL 561
 Db 586 ALDNLGQTSLHRAAYCGHLOTCRLLLSYSGDPNIIISLQGTALQMGNNVQQLQEQEISL 645
 QY 562 GNSEADRLLEAAKAGDVTETVKKLCTVQSNCRDIEGRQSTPLHFAAGYNNRVSVVYLLQ 621
 Db 646 RTSDVDYRLLEASKAGDLETVKLCSQNVNCRDIEGRHSTPLHFAAGYNNRVSVVYLLH 705
 QY 622 HGADVHAKDKGGLVPLHNACSYGHEVAELLVKHGVNVNADLWKETPLHFAAAGKYEI 681
 Db 706 HGADVHAKDKGGLVPLHNACSYGHEVAELLVKHGVNVNADLWKETPLHFAAAGKYEI 765
 QY 682 CKLLQHGADPTKKNRDGNTPLDLVKDGTIDIDLLRGDAALLDAKKGCLARVKLSSP 741
 Db 766 CKLLQHGADPTKKNRDGNTPLDLVKDGTIDIDLLRGDAALLDAKKGCLARVKLCTP 825
 QY 742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVDV 801
 Db 826 ENINCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVDI 885
 QY 802 AALLIKYNCVNAATDKWAFPLHEAAQGRQLCALLAHGADPTMKNQEQOTPLDLATA 861
 Db 886 AALLIKYNCVNAATDKWAFPLHEAAQGRQLCALLAHGADPTMKNQEQOTPLDLATA 945
 QY 862 DDVSAALLTAAMPSPALPSCYKQVNLNGVRSPGATADALSQSPSSLSAASSLDNLGSG 921
 Db 946 DDIRALLIDAMPPEALPTCFKPOAT-----VVSASLISPASTPSCLSAASSIDNLGTP 998
 QY 922 FSELSSVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDFIREQITLDLV 979
 Db 999 LAELAVGASNAGDGAACERKECEVAGLDWNTISQFLKSLGLEHLRDFITEQITLDVLA 1058

```
QY 980 EMGHKELKEIGINAYGHRHKLKIGVERLISGOOGLNPYLTNTSGGTILIDLSPDDKEF 1039
Db 1059 DMGHEELKEIGINAYGHRHKLKIGVERLLGGQOQGTNPYLTFFHCVNQGTILLDLAPEDKEY 1118
QY 1040 QSVVEEMQSTVREHRDGGHAGGTFNRYNLIKQVCNKKLWERYTHRRKEVSEENHNHAN 1099
Db 1119 QSVVEEMQSTIREHRDGGNAGGTFNRYNVIKQVNVNKKLRERFCHRQKEVSEENHNHN 1178
QY 1100 ERMFLHGSPEVNAILHKGFDERHAYIGGMFGAGIYFAENSSKSNOXYVYIGGGTGCPVHK 1159
Db 1179 ERMFLHGSPPFNAILHKGFDERHAYIGGMFGAGIYFAENSSKSNOXYVYIGGGTGCPVHK 1238
QY 1160 DRSCYICHROLLEFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVYIRGE 1219
Db 1239 DRSCYICHROMLEFCRVTLGKSFLQFSMTKMAHAPPGHHSVIGRPSVNGLAYAEYVYIRGE 1298
QY 1220 QAYPEYLITYQIMRPE 1235
Db 1299 QAYPEYLITYQIMKPE 1314
```

Search completed: February 12, 2003, 03:32:21
Job time : 45.3932 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 07:42:50 : Search time 16.9573 Seconds
(without alignments)
2151.551 Million cell updates/sec

Title: US-09-843-159B-4
Perfect score: 6464
Sequence: 1 RCSARRGAAGGGAQRGARV.....AYPEYLITYQIMRPGMVDG 1240

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6053	93.6	1166	4	US-09-350-982C-5
2	5103.5	79.0	1327	4	US-09-196-387-2
3	3307	51.2	949	4	US-09-196-387-10
4	2132.5	33.0	673	4	US-09-196-387-8
5	877	13.6	1839	2	US-09-172-977-4
6	858	13.3	843	2	US-09-172-977-3
7	835	12.9	1745	2	US-09-031-485-33
8	835	12.9	1745	2	US-08-847-429A-33
9	835	12.9	1745	3	US-09-065-474-33
10	835	12.9	1745	4	US-09-557-034-33
11	629.5	9.7	1088	4	US-09-082-059-2
12	514.5	8.0	352	3	US-09-065-474-139
13	514.5	8.0	352	4	US-09-557-034-139
14	421.5	6.5	303	2	US-09-031-485-23
15	421.5	6.5	303	2	US-08-847-429A-23
16	421.5	6.5	303	3	US-09-065-474-23
17	421.5	6.5	303	4	US-09-557-034-23
18	409	6.3	302	2	US-09-031-485-38
19	409	6.3	302	2	US-08-847-429A-38
20	409	6.3	302	3	US-09-065-474-38
21	409	6.3	302	4	US-09-557-034-38
22	398.5	6.2	1719	4	US-09-605-785-378
23	398.5	6.2	1719	4	US-09-439-313-378
24	398.5	6.2	1719	4	US-09-352-616A-378
25	392	6.1	348	2	US-09-031-485-28
26	392	6.1	348	2	US-08-847-429A-28
27	392	6.1	348	3	US-09-065-474-28

28	392	6.1	348	4	US-09-557-034-28	Sequence 28, Appl
29	390.5	6.0	787	4	US-09-188-930-334	Sequence 334, App
30	345	5.3	1423	4	US-09-810-712-10	Sequence 10, Appl
31	332.5	5.1	835	4	US-09-291-839-2	Sequence 2, Appl
32	310	4.8	191	2	US-09-031-839-2	Sequence 20, Appl
33	310	4.8	191	2	US-08-847-429A-20	Sequence 20, Appl
34	310	4.8	191	3	US-09-065-474-20	Sequence 20, Appl
35	310	4.8	191	4	US-09-557-034-20	Sequence 20, Appl
36	306	4.7	422	2	US-08-484-575A-6	Sequence 6, Appl
37	306	4.7	422	3	US-08-477-459-6	Sequence 6, Appl
38	306	4.7	422	3	US-08-479-869-6	Sequence 6, Appl
39	306	4.7	422	4	US-08-486-414-6	Sequence 6, Appl
40	306	4.7	422	5	PCT-US94-01826A-6	Sequence 6, Appl
41	306	4.7	422	5	PCT-US94-02352A-6	Sequence 6, Appl
42	305	4.7	679	2	US-08-462-481-4	Sequence 4, Appl
43	305	4.7	679	2	US-08-436-771-6	Sequence 6, Appl
44	305	4.7	679	2	US-08-434-998-6	Sequence 6, Appl
45	305	4.7	679	2	US-08-487-797-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-350-982C-5
: Sequence 5, Application US/09350982C
: Patent No. 6455290
: GENERAL INFORMATION:
: APPLICANT: Berthelsen, Jens
: APPLICANT: Toma, Salvatore
: APPLICANT: Isacchi, Antonella
: TITLE OF INVENTION: Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods R
: TITLE OF INVENTION: Same
: FILE REFERENCE: PHRM-0043
: CURRENT APPLICATION NUMBER: US/09/350,982C
: CURRENT FILING DATE: 1999-07-09
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 5
: LENGTH: 1166
: TYPE: PRT
: ORGANISM: Artificial
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (1102)..(1102)
: OTHER INFORMATION: n is any nucleic acid
: NAME/KEY: misc_feature
: LOCATION: (2650)..(2650)
: OTHER INFORMATION: n is any nucleic acid
US-09-350-982C-5

Query Match 93.6%; Score 6053; DB 4; Length 1166;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1157; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY	75	MSGRRCAGGGAACASAAAEVPAARELFACRNGDVERVKRLVTPKVNSTRDTAGRKST	134
Db	1	MSGRRCAGGGAACASAAAEVPAARELFACRNGDVERVKRLVTPKVNSTRDTAGRKST	60
QY	135	PLHPAAGFGRKDVVEYLQNCANVOARDGGLIPLHNACSFGEAEVNNLLRHGADPNAR	194
Db	61	PLHLAAGFGRKDVVEYLQNCANVOARDGGLIPLHNACSFGEAEVNNLLRHGADPNAR	120
QY	195	DNWNYTLPHEAAIKGKIDVCIVLQHGAEPTIRNTDGTALDLPSPSAKAVLTGEYKDE	254
Db	121	DNWNYTLPHEAAIKGKIDVCIVLQHGAEPTIRNTDGTALDLPSPSAKAVLTGEYKDE	180
QY	255	LLSARSNGEKKMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK	314
Db	181	LLSARSNGEKKMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK	240

```
QY 315 DRGDLVPLHNACSYGHEYTELLVKGACVNAAMDLMQFTPLHEAASKNRVEVCSLLLSYG 374
Db 241 DRGDLVPLHNACSYGHEYTELLVKGACVNAAMDLMQFTPLHEAASKNRVEVCSLLLSYG 300
QY 375 ADPTLLNCINKSAIDLAPTQPKERLAYEFKGHSLLQAAAREADVTRIKKHSLSLEWVNFKH 434
Db 301 ADPTLLNCINKSAIDLAPTQPKERLAYEFKGHSLLQAAAREADVTRIKKHSLSLEWVNFKH 360
QY 435 POTHETALHCAASAPYKPKQKQCELLLRKGANINEKTKFELPLHVASEKAHNDVVEVV 494
Db 361 POTHETAXHCAASAPYKPKQKQCELLLRKGANINEKTKFELPLHVASEKAHNDVVEVV 420
QY 495 KHEAKYNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLOGFTALQMGNNVQQL 554
Db 421 KHEAKYNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLOGFTALQMGNNVQQL 480
QY 555 LQEGISLGNSEADROLLEAKAGADVETVKKLCVTQSVNCRDTEGROSTPLHFAAGYNRV 614
Db 481 LQEGISLGNSEADROLLEAKAGADVETVKKLCVTQSVNCRDTEGROSTPLHFAAGYNRV 540
QY 615 VVEYLLOHGADVHAKDKGLVPLHNACSYGHEVEAEVLLVKGAVVNVADLWKFTPLHEAA 674
Db 541 VVEYLLOHGADVHAKDKXXLVPLHNACSYGHEVEAEVLLVKGAVVNVADLWKFTPLHEAA 600
QY 675 AKGYEICKLLLOHGADPTKKNRDNPTPLDLVKDGTDTQDLRLGDAALLDAKKGCLAR 734
Db 601 AKGYEICKLLLOHGADPTKKNRDNPTPLDLVKDGTDTQDLRLGDAALLDAKKGCLAR 660
QY 735 VKKLSPDNVNCRDTOGRHSTPLHLAAGYNNLEVAEYLLLOHGADVNAQDKGGLIPLHNA 794
Db 661 VKKLSPDNVNCRDTOGRHSTPLHLAAGYNNLEVAEYLLLOHGADVNAQDKGGLIPLHNA 720
QY 795 SYGHVDVAALLIKYNACVNATDKWFTPLHEAAKQRTQCALLLAHGADPTLKNGEGOT 854
Db 721 SYGHVDVAALLIKYNACVNATDKWFTPLHEAAKQRTQCALXLAHGDPTLKNGEGOT 780
QY 855 PLDLVSADVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSSGSPSSLSAAS 914
Db 781 PLDLVSADVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSSGSPSSLSAAS 840
QY 915 LDNLGSFSLSSVSSSTEGASSLEKKEVPCVDFSIQFVRLGLEHLMDFEREQIT 974
Db 841 LDNLGSFSLSSVSSSTEGASSLEKKEVPCVDFSIQFVRLGLEHLMDFEREQIT 900
QY 975 LDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGQGLNPYLLTNTSGSGTILDLSP 1034
Db 901 LDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGQGLNPYLLTNTSGSGTILDLSP 960
QY 1035 DKEFOSVEEEMOSTVREHRDGGHAGGIENRNNILKIQVCNKKLWERTYTHRKEYSEEN 1094
Db 961 DKEFOSVEEEMOSTVREHRDGGHAGGIENRNNILKIQVCNKKLWERTYTHRKEYSEEN 1020
QY 1095 HNHANERMLFHGSPFFNAILHKGFDERHAYIGMFCAGIYFAENSSKSNQYVYIGGGTG 1154
Db 1021 HNHANERMLFHGSPFFNAILHKGFDERHAYIGMFCAGIYFAENSSKSNQYVYIGGGTG 1080
QY 1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGPHSHVTCRPSVNGLALAEV 1214
Db 1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGPHSHVTCRPSVNGLALAEV 1140
QY 1215 IYRGEQAYPEYLITYQIMRPEGVMDG 1240
Db 1141 IYRGEQAYPEYLITYQIMRPEGVMDG 1166
```

RESULT 2

US-09-196-387-2

; Sequence 2, Application US/09196387

; Patent No. 6277613

; GENERAL INFORMATION:

; APPLICANT: de Lange, Titia

; APPLICANT: Smith, Susan

; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRP1 AND METHODS

```
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,387
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/095,225
; FILING DATE: June 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1327 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; US-09-196-387-2
```

Query Match

79.0%; Score 5103.5; DB 4; Length 1327;

Best Local Similarity 79.2%; Pred. No. 0;

Matches 963; Conservative 110; Mismatches 128; Indels 15; Gaps 3;

```
QY 22 AAHTGAPDPTVTSQAARALSSPGGLALLLAGPGLLRLLALLLAVAAARIMSGRCA 81
Db 112 SAAGVAPNPAGSGNNSSPSSPTSSSSSSPSG-----SSLAESPEAAGVSSTAPL 165
QY 82 GGAACASAAAEEAVEPAARELFEACRNGDVERVKRLVTEKVNRODTAGKSTPLHFAAG 141
Db 166 GGAAGGPGTGVPAVGALRELLACRNGDVSVRKRLVDAANVNAKDMAGRKSPHFAAG 225
QY 142 FGRKDVVEYLLQGANVQARDGGLIPLHNACSFHAEVYNLLLRHGADPNARDNNYTP 201
Db 226 FGRKDVVEYLLQGANVHARDGGLIPLHNACSFHAEVYNLLLRHGADPNARDNNYTP 285
QY 202 LHEAAIKGIDVCIVLLQGAETPIRNTDGTALDLADPSAKAVLTGEYKDELLESARS 261
Db 286 LHEAAIKGIDVCIVLLQGAETPIRNTDGTALDLADPSAKAVLTGEYKDELLESARS 345
QY 262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAGYNRKIVQLLOHCAADVHAKDGLVP 321
Db 346 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAGYNRKIVQLLOHCAADVHAKDGLVP 405
QY 322 LHNACSYGHEYTELLVKGACVNAAMDLMQFTPLHEAASKNRVEVCSLLLSYGADPTLN 381
Db 406 LHNACSYGHEYTELLVKGACVNAAMDLMQFTPLHEAASKNRVEVCSLLLSYGADPTLN 465
QY 382 CHNKSALDAPTQPKERLAYEFKGHSLLQAAAREADVTRIKKHSLSLEWVNFKHQTHETA 441
Db 466 CHKSADVMAPTPELRLTYEFKGHSLLQAAAREADVTRIKKHSLSLEWVNFKHQTHETA 525
QY 442 LHCAASPYKPKQKQICELLLRKGANINEKTKFELPLHVASEKAHNDVVEVYVYHKAQVN 501
```

Db	526	LHC	AVASLHPKRVQVTELLLRKGANVYKNDQMTPLHVAA	GRAHNDVNEVLHKHGAKMN	585
Qy	502	ALD	NLGOTSUHRAAAYCHLQTCRLLSYSGCDPNIIIS	LOGFTALQMGNEVNOQLQEGTSL	561
Db	586	ALD	TGQATLHAALAGHLQTCRLLSYSGDPSIIS	LOGFTAAQMGNEAVQOILSESTPI	645
Qy	562	GNSE	ADROLLEAAKAGDVEYVKKLTVQSVNCRDTEG	QSGTPLHPFAAGYNRVSVVEYLLQ	621
Db	646	RTSD	VDYRLLEASKAGDLETQKLCSSQNVNCRDLEGR	HSHTPLHPFAAGYNRVSVVEYLLH	705
Qy	622	HGAD	VHAKDKGGLVPLHNAACSYGHVEABELLVKH	GAVNVADLNKFTPLHEAAAAGKYEI	681
Db	706	HGAD	VHAKDKGGLVPLHNAACSYGHVEABELLVHR	GASVNVADLNKFTPLHEAAAAGKYEI	765
Qy	682	CKLL	QHOGADPTKKNRDGNTPDLVKGDDTDIQD	LLRGDAALLDAAKKGLCARVKKLSSP	741
Db	766	CKLL	LKHGADPTKKNRDGNTPDLVKEGDDTDIQD	LLKGDGDAALLDAAKKGLARVQKLCTP	825
Qy	742	DNVC	RDTOGRHSTPLHLAAGYNNI	EVARYLLOHGADVNAQDKGLIPLPHNAASYGHVDV	801
Db	826	ENTN	CRDTOGRNSTPLHLAAGYNNI	EVAREYLLLEHGADVNAQDKGLIPLPHNAASYGHVDI	885
Qy	802	AALL	KYNACVNATDKWATPTLHEAAQKGR	TOLCALLLAHGADPTLKNQEGOTPLDLVSA	861
Db	886	AALL	KYTCVNATDKWATPTLHEAAQKGR	TOLCALLLAHGADPTMKNQEGOTPLDLATA	945
Qy	862	DVVS	ALLTAAMPSPALPSCKPQVLNVGRSP	GATADALSSGSPSSPSLSAASSLDNLSGS	921
Db	946	DDI	RALLIDAMPPEALPTCFKPKQAT	-----VVSASLISPASTPSCLSAASSIDNLGTG	998
Qy	922	FSLS	SVWSSGTEGASSLEKK--EYVGVDP	SITQFVNRLGLEHLMIDIFEREQITLDVLV	979
Db	999	LAEV	AGSADGAGAEPTKEGEVAGIDMNI	ISOFLKSLGLEHLRDIETEFETITDVL	1058
Qy	980	EMGH	KEKLGINAYGHRHKLKGV	ERLISGQOGLNPYLLTNTSGSGTILDLSPDDKEF	1039
Db	1059	DMG	HEELKEIGINAYGHRHKLKGV	ERLLGGOQTNPYLLTFCVNOQTILDLDAPEQY	1118
Qy	1040	QSVE	EMQSTVREHRDGGHAGGIFNRY	NILKIQKCNKKLWERYTHRRKEVSEENHNHAN	1099
Db	1119	QSVE	EMQSTIREHRDGGNAGGIFNRY	NIIRIQKVVKKLREFCRQKEVSEENHNHN	1178
Qy	1100	ERML	PHGSPFVNAIITHKGF	DERHAYIGGMFCAGIYFAENSSKSNQYVYGIGGTCGPVHK	1159
Db	1179	ERML	PHGSPFINAIITHKGF	DERHAYIGGMFCAGIYFAENSSKSNQYVYGIGGTCGPETHK	1238
Qy	1160	DRSC	YCHROLLFCRVTLCKSKFLQF	SAMKMAHSPGHHSVTGRPSVNCALAEVYVYRGE	1219
Db	1239	DRSC	YCHRQMLFCRVTLCKSKFLQF	STMKMAHPPGHHSVTGRPSVNCALAEVYVYRGE	1298
Qy	1220	QAYP	EYLITYIQMRPE	1235	
Db	1299	QAYP	EYLITYIQMKPE	1314	

```

RESULT 3
US-09-196-387-10
: Sequence 10, Application US/09196387
: Patent No. 6277613
: GENERAL INFORMATION:
: APPLICANT: de Lange, Titia
: APPLICANT: Smith, Susan
: TITLE OF INVENTION: A PROTEIN THAT BINDS TO TR1 AND METHODS
: TITLE OF INVENTION: OF USE THEREOF
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauber & Jackson
: STREET: 411 Hackensack Avenue, 4th Floor
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER READABLE FORM:

```

Db 706 HGADVHAKDKGGLVPLHNACSYGHVEAEVLLVRHGASVNVADLWKFPTPLHEAAAKGKYEI 765
QY 682 CKLLQHGADPTKKNRDNTPDLVYKGDGTDQDILLRGDAALLDAAKCKCLARVKKLSSP 741
Db 766 CKLLKHGADPTKKNRDNTPDLVYKGDGTDQDILLRGDAALLDAAKCKCLARVQKLCPT 825
QY 742 DAVNCRDGTGRSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH--V 799
Db 826 ENINCRDGTGRSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGGCLA 885
QY 800 DVAALLIKYNACVNAIDKWAFTPLHEAAQGRGTQCALLLAHGADPTLKNQEGOTPL 856
Db 886 RVQKCTPENINCRDGTGRSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPL 942

RESULT 4
US-09-196-387-8
; Sequence 8, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRP1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,387
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/095,225
; FILING DATE: June 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 673 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-196-387-8

Query Match 33.0%; Score 2132.5; DB 4; Length 673;
Best Local Similarity 73.3%; Pred. No. 7.6e-177;
Matches 417; Conservative 54; Mismatches 91; Indels 7; Gaps 2;

QY 22 AANGTAPDPTAGSQARALSSPGGLALLAGPGLILRLALLLAVAAARIMSGRCA 81
Db 112 SAAGVAPNPAGSGNSPSSSSPTSSSSPSPG-----SSLAESPEAAGVSTAPL 165
QY 82 GGCAACASAAAEVAPARELFACRNGDVERVKRLVTPPEKVNRSRTAGRKSTPLHPAAG 141
Db 166 GPAGAGPCTGPAVSGALRELLACRNGDVSRYKRLVDAANVNAKMDAGRKSSPLHPAAG 225
QY 142 FGRKDVVEYLLQNGANVQARDGGLIPLHNACSFHGAEVVNNLLRHGADPNARDNNWYTP 201

Db 226 FGRKDVVEYLLQNGANVHARDGGLIPLHNACSFHGAEVVNNLLRHGADPNARDNNWYTP 285
QY 202 LHEAAIKGIDVICIVILQHGAEPTIRNTDGRDALDADPSAKAVLTGEYKKDELLESARS 261
Db 286 LHEAAIKGIDVICIVILQHGADPNIRNTDGRSALDADPSAKAVLTGEYKKDELLESARS 345
QY 262 GNEEKHMAILLPLNVNCHASDGRKSTPLHLAAGYNNRVKIVQVLLQHGADVHAKDKGDLVP 321
Db 346 GNEEKHMAILLPLNVNCHASDGRKSTPLHLAAGYNNRVKIVQVLLQHGADVHAKDKGDLVP 405
QY 322 LHNACSYGHEYTELVKKGACVNAIDKWAFTPLHNAASYGH--V 381
Db 406 LHNACSYGHEYTELVKKGACVNAIDKWAFTPLHNAASYGH--V 465
QY 382 CHNKSALDAPTPOLKERLAYEFKGHSLLOAAAREADVTRIKKHLSLVNNVFKHPQTHETA 441
Db 466 CHKSALDAPTPOLKERLAYEFKGHSLLOAAAREADVTRIKKHLSLVNNVFKHPQTHETA 525
QY 442 LHCAASPYPKRQICICELLRLKGANINEKTEKFTPLHVAASEKAHNDVVEVYVYKHEAKVN 501
Db 526 LHCAVASLHPKRQVTELLRLKGANVNEKNKDFMTPLHVAASEKAHNDVVEVYVYKHEAKVN 585
QY 502 ALDNLQOTSLHRAAYCCHLQTCRLLLSYCGDPMIISLQGFALOMGNVQVLLQEGISL 561
Db 586 ALDNLQOTSLHRAAYCCHLQTCRLLLSYCGDPMIISLQGFALOMGNVQVLLQEGISL 645
QY 562 GNSEADROLLEAAKAGDVETVKKILCTVQS 590
Db 646 DFIISLQGFALOMGN--EAVQVLLSGHS 673

RESULT 5
US-09-172-977-4
; Sequence 4, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g29491
US-09-172-977-4

Query Match 13.6%; Score 877; DB 2; Length 1839;
Best Local Similarity 29.7%; Pred. No. 1.1e-66;
Matches 293; Conservative 137; Mismatches 402; Indels 156; Gaps 23;

QY 105 ACRNGDVER-VKRLVTPPEKVNRSRTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARD 163
Db 38 AARAGNIDKVEYKKGIDINTCNQGLNA--LHAAKEGHVGLVQELLGRGSSVDSATK 95
QY 164 GGLTPLHNACSFHGAEVVNNLLRHGADPNARDNNWYTPLEHAAIKGIDVICIVILQHGAE 223
Db 96 KGNTALHIALSLAQAEVVKVVLKREGANINQASQNGFTPLYMAQENHIDVVKVLLNGAN 155
QY 224 PTIRNTDGR-----ALDADPSAKAVLTGEYKKDEL-----LESARSGNEKMMALLPLN 275
Db 156 QSTATEDGFTPLAVALQOQHNAVAIILENDTKGKVRPLPALHIAARKDDTKSAALLQ-- 213
QY 276 VNCHASDGRKS-----TPHLAAGYNNRVKIVQVLLQHGADVHAKDKGDLVPLHN 324


```
Db 214 -NDHNADVQSKMVMNRTTESGFTPLHIAAHYGNVNVATLLNRCGAADVFTARNGITPLHV 272
QY 325 ACSYGHYEVTPELLVKGACVNMALWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db 273 ASKRGNTNMVKLLDRGGQIDAKTRDGLTPLHCARSGHQDVVELLERGA-PLIARTKN 331
QY 385 KSAIDLAPTOLKERLAYEFKGSLLQAAREADVTRIKKHLSEVMNFKHP-----QTHE 439
Db 332 -----GLSPLHMAAQGDHVECVKHL-----LQHKAPVDVDTLDYL 366
QY 440 TALHCAASPYPRKQICELLRRKGANINEKTEFLTPHVAASKAHNDVVEVVKHBEAK 499
Db 367 TALHVA--HCGHYRVTKLLDKRANPARNALNGFTPLHIAACKNRKIKVMELLYKYGAS 423
QY 500 VNALDNLGQTSLHRAAYCGHLCOTCRLLLSYCCDPNIIISLOGFTALOMGN-----ENVQOL 554
Db 424 IQAITESGLTPHVAAPFNGHLNIVLLLLONGASPDVNIIRGETALHMAARAGQVEVVRCL 483
QY 555 LQEGI-----SLGNSEADROLLE-----AAKAGDV 579
Db 484 LRNGALVDARAREBOTPLHIASRLGKTEIVOLLLOHMAHPDAATTNGYTPHISAREQV 543
QY 580 ETYVKKLTQVOSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHN 639
Db 544 DVASVLEAGAAHSLATK-RGFTPLHVAAYKYSLDVAKLLQRRRAADSAGKNGLTPLHV 602
QY 640 ACSYGHYEVAELLVKGACVNVADLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 699
Db 603 AAHYDNQKVALLEKASPHATAKNCTYTPHIAAKNQMQIASTLLNYGAETNIVTKQG 662
QY 700 NTPLDLV-KGDDTDIOLDRGDAALDAACKGLCLARYKLLSSPDNVNCRDQGRHSTPLH 758
Db 663 VTPLHLASQEGHDMVTLLDKGANIHMSTKSL-----TSLH 700
QY 759 LAAGYNNEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKW 818
Db 701 LAAGKVNVDIILTHGADODDHTKLGTYPLIVACHYGNVMKVNFLLLKOGANVNAKTKN 760
QY 819 AFTPLHEAAQKGTOLCALLLHAGADPTLKNQEGQTPDL-----VSADDVSALLTAAM 872
Db 761 GYTPPLHQAQOQTHIINVLLQHGAKPNATTANGNTALATAKRLGISVVDTLKVYTEEV 820
QY 873 PPSALPSCYKQVINGVRSPGATADALSSGSPSSLSAASSIDNLSGSFSEL-----925
Db 821 TTTTITITEK-----HKLNPETMTEVL-----DVSDEEGDDTMTGCGEYLRPEDLK 868
QY 926 -----SSVSVSSGTEGASSLEKKEVPGVDFEITQFV--RNLGLEHLMDFEREQITLDVLV 979
Db 869 ELGDDSLPSSQFLDGMNLYLSLEGGRSDSLRFSFSDRSHTLSHAS--YLURDSAVMDDSV 926
QY 980 EMGHKELKEIGINAYGHRHKLKIGVERL 1007
Db 927 VIPSHQVSTLAKAERNYSYRLSWGTENL 954
```

RESULT 6

```
US-09-172-977-3
; Sequence 3, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Neil J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 843
; TYPE: PR
```

```
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: g1841966
US-09-172-977-3
```

```
Query Match 13.3%; Score 858; DB 2; Length 843;
Best Local Similarity 29.6%; Pred. No. 1.3e-65;
```

```
Matches 275; Conservative 148; Mismatches 367; Indels 140; Gaps 23;
```

```
QY 105 ACRNGDVER-VKRLVTPKVNRSRTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDD 163
Db 3 AARAGNLDKVVLEYLKGIDINTCQNGUNA--LHLAAKEGHVGLVQELLGRSSVDSATK 60
QY 164 GGLIPLHNACSFHAEVYVNLRLRHGADPNARNWNVYTPLHEAAIKGKIDVICVILLOHGA 223
Db 61 KGNTALHIAASLAGAEVYVNLVKEGANINAOQNGFTPLYMAAQENHIDVYVYLLLENGAN 120
QY 224 PTIRNTDGT-----ALDLADPSAKAVLTGEYKQDEL-----LESARSGNEEKMMALLTPLN 275
Db 121 QSTATDGTPLVALOQHNQAVAILLENDTKGVRLPALHIAARKDDTTSAALLQ-- 178
QY 276 VNCHASGRKS-----TPLHLAAGYNRVKIVOLLQHGADVHAKDKGDLVPLHN 324
Db 179 -NDHNADVQSKMVMNRTTESGFTPLHIAAHYGNVNVATLLNRCGAADVFTARNGITPLHV 237
QY 325 ACSYGHYEVTPELLVKGACVNMALWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db 238 ASKRGNTNMVKLLDRGGQIDAKTRDGLTPLHCAARSGHQDVVELLERGA-PLIARTKN 296
QY 385 KSAIDLAPTOLKERLAYEFKGSLLQAAREADVTRIKKHLSEVMNFKHP-----QTHE 439
Db 297 -----GLSPLHMAAQGDHVECVKHL-----LQHKAPVDVDTLDYL 331
QY 440 TALHCAASPYPRKQICELLRRKGANINEKTEFLTPHVAASKAHNDVVEVVKHBEAK 499
Db 332 TALHVA--HCGHYRVTKLLDKRANPARNALNGFTPLHIAACKNRKIKVMELLYKYGAY 388
QY 500 VNALDNLGQTSLHRAAYCGHLCOTCRLLLSYCCDPNIIISLOGFTALOMGNVQQLQBEI 559
Db 389 IQAITESGLTPHVAAPFNGHLNIVLLLLONGASPDVNIIRGETALHM-----435
QY 560 SLGNSEADROLLEAAKAGDVETVKKLTQVOSVNCRDIEGR-QSTPLHFAAGYNRVSVVEY 618
Db 436 -----AARAGEVEVVR--CLLRNGALVDARAREEQTPHIASRLGTEIVOL 480
QY 619 LQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKGACVNVADLWQFTPLHEAASKGK 678
Db 481 LLOHMAHPDAATTNGYTPHISAREGOVDVASVLEAGAAHSLATKKGFTPLHVAAYKGS 540
QY 679 YEICKLLOHGAADPTKKNRQDNTPLDLVKDGDITDQLDRGDAALLDAACKGLARVKKL 738
Db 541 LDYAKLLQRRRAADSAGKNGLTPLHVAAYDN-----QKVALLLLEK 584
QY 739 SSPDNVNCRDQGRHSTPLHIAAGYNNEVAEYLLQHGADVNAQDKGLIPLHNAASYGH 798
Db 585 ASP-----HATAKNGYTPHIAAKKNQMQIASTLLNYGAETNIVTKGVTPHLSAQEGH 639
QY 799 VDVAALLIKYNACVNAATDKWAFTPHIAAAGKGTOLCALLLHAGADPTLKNQEGQTPDL 858
Db 640 TDMVTLVLEKGANIHMSTKSGLTSLHIAAEEKVNVAADILTKHGADQDADATKLGTYTPLL 699
QY 859 V-----SADDVSALLTAAMPESA-LPCYK-----QVLNGVRSFGATDALSSGP 903
Db 700 ACHYGNVKNVNFLLKOGANVNAKTNGYTPHQAQOQTHIINVLLQHGAKPNATTA-- 757
QY 904 SSPSSLSAASSLDNLSGFSFSELSSVSSSGTEGASSLEKKEVPGVDFEITQFVRNLGLEH 963
Db 758 NGNTALAIARLGYI--SVVDTLKVVTTEVTTTTTITEKHKLNAPEMTTE-----806
QY 964 LMDIFERE---QITLDVLVEMGHKELKEIG 990
Db 807 VLDVDSDEEGDDTVTGDGGEYLRPELDKELG 836
```

RESULT 7

US-09-031-485-33

; Sequence 33, Application US/09031485

; Patent No. 5824306

; GENERAL INFORMATION:

; APPLICANT: Tang, Liang

; APPLICANT: Blehm, E. Scot

; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND

; TITLE OF INVENTION: USES THEREOF

; NUMBER OF SEQUENCES: 85

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carol Talkington Verser, Ph.D.

; ADDRESS: Hesk Corporation

; STREET: 1825 Sharp Point Drive

; CITY: Fort Collins

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80525

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Windows 95

; SOFTWARE: WordPerfect for Windows, Version 7.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/031.485

; FILING DATE:

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/847.429

; FILING DATE: 24-APR-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Verser, Carol Talkington

; REGISTRATION NUMBER: 37,459

; REFERENCE/DOCKET NUMBER: HW-5

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 970/493-7272

; TELEFAX: 970/484-9505

; INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1745 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-031-485-33

Query Match 12.9%; Score 835; DB 2; Length 1745;

Best Local Similarity 25.7%; Pred. No. 4.5e-63;

Matches 314; Conservative 154; Mismatches 452; Indels 304; Gaps 37;

```
QY 34 GSOAARALSASSPGGLALLAGPGLLLRLALLLAVAAARIMSG---RRC-AGGGAACAS 89
Db 35 GESSASFRAARAGNLD-----RVLELL-----RSGTDINTCANGNLNHL 76
QY 90 AAAAEVAPAAARELFACRNGDVERVKRLVTPKVNRSRDITAGRK-STPLHFAAGFGKDVV 148
Db 77 ASKEGHVRELLK--RKADV-----DAATRKGNLTALHSLAQOELIV 119
QY 149 EYLLONGANVQARDGGGLIPLHNACSFHAEVYVNLRLRHGADPNARDWNYTPLHEAAIK 208
Db 120 TVLVENGANVYQSLNGFTFLYMAAQENHESVRYLLAHNAQALSTEDGFTPLAVALQQ 179
QY 209 GKIDVCIVLHGAETIRNTDGRALDIA-----DPSAKAVLTGETYKKDELLES----- 258
Db 180 GHDRVAVULENDTRGKVR-----LPALHTAAKDDTKAATLLQNEHNSDVTSKSGFTPL 235
QY 259 ---ARSGNEEKMMALL-----TPLNV-----NCHAS 281
Db 236 HIAAHYGNENVAQLLLEKCANVYQARHNISPLHVATKWRTKMVSVLLLAHGAVIDCRTR 295
QY 282 DGRKSTPLHAGYKRVKQLVQLLHOGADVHAKDKDGLVPLHNACSY----- 328
```

```

; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heskia Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-847-429A-33

Query Match 12.9%; Score 835; DB 2; Length 1745;
Best Local Similarity 25.7%; Pred. No. 4.5e-63;
Matches 314; Conservative 154; Mismatches 452; Indels 304; Gaps 37;

QY 34 GSOARALSSPGGLALLAGPGLLRLLIALLLAVAAARIMSG---RRC-AGGGAACAS 89
DB 35 GESSASFLRAARAGND-----RVLELL-----RSGTDINTCANGNALHL 76
QY 90 AAEAVEPAARELFEACRGDVERVRLVTPKVNRSRDYAGRK-STPLHFAAGFGKDDVV 148
DB 77 ASKEGHEVYVRELLK--RKADV-----DAATRGNTALHIASLAGOELIV 119
QY 149 EYLLONGANVQARDGGLPLHNACSFHGAEVVNLRLHCGADPNARDNNWYTPHAAIK 208
DB 120 TVLVENGANVNSLNGFTPLMAAENHESVYVRYLLAHNANQALSTEDGFTPLAVALQQ 179
QY 209 GKIDVCIVLQHGAEPTIRNTDGRALDLA-----DPSAKAVITGEYKKDELLES----- 258
DB 180 GHDRVVAVLLENDTRGKVR---LPALHIAAKKDDTKAATLLQLQNEHNSDVTSGFTPL 235
QY 259 ---ARGNEKMMALL-----TPLNV-----NCHAS 281
DB 236 HIAAHTGNEVNAVLLEKGANVYQARHNISPLHVAATKMGRTNMVSLLLAHGAVIDCTR 295
QY 282 DGRKSTPLHLAAGYNNVKIVOLLQHCADVHAKDGLVPLHNACSY----- 328
DB 296 D--LLTPLHCASRGHDQVVDLLEKAPISAKTKNGLAPLHMAAQVDDVTDYDYLPLHV 353
QY 329 ----GHYEVTELLVKHCACVANDMLQWFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
DB 354 AAHCGHVRVAKLLDRNADPNARALNGFTPLHIAACKNRIKIVELLKYHA----- 404
QY 385 KSAIDLAPTPOLKRLAYEPKGH-----SLQAARADVTIRIKKHLKSLWMNFKHPOTHE 439
DB 405 --AIEATTESGLSPLHVAAPMGAINIVIYLLQOGANADVATV-----GE 447
QY 440 TALHCAASPYPRKQICELLKRGANINEKTEFTPLHVAASEKAHNDVVEVVVHEAK 499
DB 448 TPLHLARA---NOTDIVRVLRNAGQVDAARAELQTPPLHIASRLGNTDIVILLQANAS 504

; 500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIISLOGFTALQMGNEVQOLLQEGI 559
; 505 PNAATRDLYTPLHIAAKEGOEVAAILMDHGTDKTLTKKGFTPLHL----- 551
; 560 SLGNSEADROLLEAAKAGOVETVKKLCCTVOSVNCRDIEGR-QSTPLHFAAGYNNRVSVVEY 618
; 552 -----AAKYGNLPVAKSL--LERGTPVDIEGKNQVTPPLHVAARYNNNDKVALL 596
; 619 LLOHGADVHAKDKGGLVPLHNACSYGHYEVAELLVHKGAVNVNADLWKEFTPLHEAAAKGK 678
; 597 LLENGASAHAAKNGYTPPLHIAAKNQMDIATLLHYKANANAESKAGFTPLHIAAEGH 656
; 679 YEICKLLLOHGADPTKKNRDGNTPLDLVKDGDIDODLLRGDAALLDAKKGCLARYKKL 738
; 657 REMAALLIENGAKVGAQRNG----- 677
; 739 SSPDNVNCRTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNADQKGLLPLHNAASYGH 798
; 678 -----LTPMHLCAQEDRVSAEELVKENAAIDPKTKAGYTPPLHVACHFGQ 722
; 799 VDVAALLIKYNACVNATDKWAFTPLHEAAOKGRTQLCALLLAHGADPTLKNQEGOTPLDL 858
; 723 INMVRFLIEHCAVSVITRASYPPLHQAQOQHNSVYVRYLLEHCASPNVHTSTGTPLSI 782
; 859 VS-----ADVSALLT-----AAMPPSALPCYKQVNLNGVRSPGATADAL---SSGPSS 905
; 783 AERLGYVSVVEALKTITETVITETITVTEERYKPKQ-----NPEAMNETMFSDEDEGE 836
; 906 PSSLSRASSLDNLGSGSELSVVSSSG-----TEGASLEKEKVGCVDFSTQFVRNL 959
; 837 DNOITANAHADFESLTK--GLHDSGTGVHLIHATEPTLS-RSPEVEGTGDDLDLIRKA 893
; 960 GLHCL-----MDIFEREQITDLVLVEMGHKELKEIGINAYGH-----RHKLKIGVER 1006
; 894 QHEPITAMADPSLDASLPDNVYI-MRTTMQPSFLISFEMVDARGAMRGCRH---SGVRI 949
; 1007 LISGQOGLNP-YLTNLTSGSGTI-----LIDLSPDKKEFOSV----- 1042
; 950 IIPPRKAPOTRVTCRYLKGDKLAHPPPLSEGEALASRILEMAPHGAKFLGVPVILEVPHF 1009
; 1043 ----EEMOSTVREHRDGHAGGIFNRYNLIKTKVCNKKLWERYTHRRKEVSEENHHA 1098
; 1010 ASLRGRERETVILRSSDQGH---WKEHQLEATEDAVQEVNLNESFDAELSOLDLHTSR 1065
; 1099 NERMLFHGSPFVNAIHKGFDERH 1122
; 1066 ITRILNDFPMYFAVTVRVQRQEVH 1089

RESULT 9
US-09-065-474-33
; Sequence 33, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heskia Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
```



```
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 139:
US-09-557-034-139

Query Match      8.0%; Score 514.5; DB 4; Length 352;
Best Local Similarity 23.7%; Pred. No. 2.4e-36;
Matches 137; Conservative 64; Mismatches 139; Indels 237; Gaps 4;

QY 131 RKSTPLHFAAGFKDKVVEYLLQNGANVOARDGGGLIPLHNACSFHAEVNVNLLRHGAD 190
Db 10 RGETPLHARANQTDIVRVLVNGAQVDAARELQTPHLIASRLGNTDIVILLQANAS 69
QY 191 PNARDNNYTPLEHAAIKGKIDVICVILLQHGAEPTIRNTDRTALDLADPSAKAVLTGEY 250
Db 70 PNAATRDLYTPLHAAKEGQEEVAAILMDHCTKTLTKKG----- 110
QY 251 KDELLESARSNGNEEKMMALLTPLNVNCHASDGRKSTPLHLAGYNRVKIVQVLLQHGAD 310
Db 111 -----FTPLHAAKYGNLPVAKSLLEGRTP 135
QY 311 VHAQKGDVLPVHNACSYGHYEVTELVKVGACVNMADLWQFTPLHEAASKNRVEVCSLL 370
Db 136 VDIEGKNQVTPLVAAHYNDKVALLLLENGASAHAAKNGYTPPLHIAAKNQMDIASTL 195
QY 371 LSYGADPTLLNCHNKSALDAPTQPKERLAYEPKHSLLQAAAREADVTRIKHLSLEMY 430
Db 196 LHY----- 198
QY 431 NFXHPQTHETALHCAAAPYKPKQICELLRLKGANINEKTEKLTPLHVASEKAHNDVV 490
Db 199 -----KANAAESKAGTTPHLHAAQEGHREMA 225
QY 491 EVVYKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIISLQGFALQMGNE 550
Db 226 ALLIENGAKVGA----- 237
QY 551 VQQLQEGISLGNSEADRLLEAAKAGDVETVKKLCTVQSVNCRDIEGROSTPLHFAAGY 610
Db 238 -----QARNG-----LTPMHLCAQE 252
QY 611 NRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEVLLVKGAVNVNADLWKTPL 670
Db 253 DRVSAEELVKNRAIDPKTKAGYTPLVACHFGQINWVFLIEHGARVSVITRASYTPL 312
QY 671 HERAAKGYEICKLLQHGADPTKKNRDGNTPDLVK 707
Db 313 HQAAQQGHNSVRYLLEHGASPNVHTSTGTPTLSIAE 349

RESULT 14
US-09-031-485-23
; Sequence 23, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
```

```
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-031-485-23

Query Match      6.5%; Score 421.5; DB 2; Length 303;
Best Local Similarity 30.4%; Pred. No. 2.3e-28;
Matches 119; Conservative 52; Mismatches 125; Indels 95; Gaps 8;

QY 467 INEKTKEFLTPHVASEKAHNDVVEVVKHAKVNALDNLGQTSLHRAAYCGHLQTCRL 526
Db 1 VDDVTVDYLTPLHV-----AAHCGHVRVAKLL 27
QY 527 LSYGCDPNIIISLQGFALQMGNEVNVQQLQEGISLGNSEADRLLEAAKAGDVETVKKLC 586
Db 28 LDRNADPNARALNGFTPLHI-----ACKNRIKIVELL 61
QY 587 TVQSVNCRDIEGRQS---TPLHEAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSY 643
Db 62 KYHAA---TEATTESGLSPLHVAFAFMCAINIVYLLQGANADVATYRGETPLHLAARA 117
QY 644 GHYEVAELLVKHGVNVNADLWKTPLHEAAAKGYEICKLLQHGADPTKKNRDGNTP 703
Db 118 NOTDIVRVLVNGAQVDAARELQTPHLIASRLGNTDIVILLQANASPNAAATRDLYTPL 177
QY 704 DL-VKDGDTDIDLLRG---DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHL 759
Db 178 HIAAKEGQEEVAAILMDHGTCTLL--TKKG-----FTPLHL 212
QY 760 AAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVVAALLIKYNACVNATDKWA 819
Db 213 AAKYGNLPVAKSLLEGRTPVDIEGKNQVTPLVHAAHYNDKVALLLLENGASAHAAKNG 272
QY 820 FTPLHEAAKGRTOICALLLAHGADPTLKNQ 850
Db 273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303

RESULT 15
US-08-847-429A-23
; Sequence 23, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
```


Search completed: February 12, 2003, 12:27:27
Job time : 25.9573 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 03:33:09 ; Search time 25.4359 Seconds
(without alignments)
1245.508 Million cell updates/sec

Title: US-09-843-159B-4
Perfect score: 6464
Sequence: 1 RGSARRGAAGGCGAQRGARV.....AYPEYLITYQIMRPGMVDG 1240

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5602	86.7	1074	US-09-509-196A-2	Sequence 2, Appli
2	5103.5	79.0	1327	US-09-841-835-2	Sequence 2, Appli
3	3938	60.9	802	US-09-964-899-41	Sequence 41, Appl
4	3307	51.2	949	US-09-841-835-10	Sequence 10, Appl
5	2132.5	33.0	673	US-09-841-835-8	Sequence 8, Appl
6	815.5	12.6	1724	US-09-964-899-43	Sequence 43, Appl
7	401	6.2	426	US-09-908-711-70	Sequence 70, Appl
8	398.5	6.2	1719	US-10-012-896-378	Sequence 378, App
9	398.5	6.2	1719	US-09-895-793-378	Sequence 378, App
10	398.5	6.2	1719	US-09-895-814-378	Sequence 378, App
11	398.5	6.2	1719	US-09-759-143-378	Sequence 378, App
12	398.5	6.2	1719	US-09-780-669-378	Sequence 378, App
13	398.5	6.2	1719	US-09-822-827-378	Sequence 378, App
14	392	6.1	740	US-09-835-788A-12	Sequence 12, Appl
15	332.5	5.1	835	US-09-947-199-2	Sequence 2, Appli
16	324.5	5.0	835	US-09-947-199-8	Sequence 8, Appli
17	312.5	4.8	599	US-09-735-368-2	Sequence 2, Appli
18	297	4.6	551	US-09-835-788A-17	Sequence 17, Appl
19	281.5	4.4	285	US-09-835-788A-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1

US-09-509-196A-2
; Sequence 2, Application US/09509196A
; Patent No. US20020037582A1
; GENERAL INFORMATION:
; APPLICANT: DALY, Roger J.
; TITLE OF INVENTION: A Potential Effector for the Grb7 Family of Signalling
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: 1871-129
; CURRENT APPLICATION NUMBER: US/09/509,196A
; CURRENT FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: P09388
; PRIOR FILING DATE: 1997-09-23
; PRIOR APPLICATION NUMBER: PCT AU98/00795
; PRIOR FILING DATE: 1998-09-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1074
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-509-196A-2

Query Match	86.7%	Score	5602	DB	10	Length	1074		
Best Local Similarity	99.4%	Pred.	No. 0						
Matches	1068	Conservative	2	Mismatches	4	Indels	0	Gaps	0
Qy	167	1PLHNACSFHAEVYVNLRLRHGADPNARDNNWYTPLHEAAIKGKIDVCTIVLLQHGAEPTI	226						
Db	1	1PLHNACSFHAEVYVNLRLRHGADPNARDNNWYTPLHEAAIKGKIDVCTIVLLQHGAEPTI	60						
Qy	227	RNTDGRTRALDPSAKAVLTGEYKKDELLESARGNEKMMALLTPLNVNCHASDGRKS	286						
Db	61	RNTDGRTRALDPSAKAVLTGEYKKDELLESARGNEKMMALLTPLNVNCHASDGRKS	120						
Qy	287	TPHLHAACYNRVKIVOLLQHCADVHAKDGLVPLHNACSYGHVEYVTELLVKGACVNA	346						
Db	121	TPHLHAACYNRVKIVOLLQHCADVHAKDGLVPLHNACSYGHVEYVTELLVKGACVNA	180						
Qy	347	MDLQWFTPLHEAASKNRVEVCSLLSYGADPTLLCHNKSATDLAPTOLKRLAYEFGK	406						

Db 181 MDLWQFTPLHGAASKNRVEVCSLLSYGADPTLLNCKNKAIDLAPTQPKERLAYEFGK 240
QY 407 HSLQAAAREADVTRIKKHLSELMVNFKHPQTHETALHCAAAASPPKPKQICELLRRKAN 456
Db 241 HSLQAAAREADVTRIKKHLSELMVNFKHPQTHETALHCAAAASPPKPKQICELLRRKAN 300
QY 467 INEKTKEFLPLHVASKAHNDVVVVVVKHEAKVNDLNGQSLHRAAYCGHLQTCRLL 526
Db 301 INEKTKEFLPLHVASKAHNDVVVVVVKHEAKVNDLNGQSLHRAAYCGHLQTCRLL 360
QY 527 LSYGCDPNIISLOGFTALOMGNENVOOLQEGISLGNSEADROLLLEAAKAGDVETVKKLC 586
Db 361 LSYGCDPNIISLOGFTALOMGNENVOOLQEGISLGNSEADROLLLEAAKAGDVETVKKLC 420
QY 587 TVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 646
Db 421 TVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 480
QY 647 EVAELVKGAVNVADLWKFTPLHFAAGYKYEICKLLQHGADPTKKNRDNTPDLV 706
Db 481 EVAELVKGAVNVADLWKFTPLHFAAGYKYEICKLLQHGADPTKKNRDNTPDLV 540
QY 707 KDGDTDIDLGRDAALLDAAKKGCLARVKLSSPDNVNCRDTQGRHSTPLHLAAGYNNL 766
Db 541 KDGDTDIDLGRDAALLDAAKKGCLARVKLSSPDNVNCRDTQGRHSTPLHLAAGYNNL 600
QY 767 EVAELVKGADYNAQDKGLIPLHNAASYGHVDVAALLIKYNACYNATDKWAFPLHEA 826
Db 601 EVAELVKGADYNAQDKGLIPLHNAASYGHVDVAALLIKYNASLNATDKWAFPLHEA 660
QY 827 AQKGRQLCALLLAHGADPTLKQEGQTPLDIVSADDSALLTAAMPSPALPSCYKQV 886
Db 661 AQKGRQLCALLLAHGADPTLKQEGQTPLDIVSADDSALLTAAMPSPALPSCYKQV 720
QY 887 NGVRSFGATADALSSGSPSSLSAASSLDNLGSFSELSSVSSVSSGTEGASSLEKKEVP 946
Db 721 NGVRSFGATADALSSGSPSSLSAASSLDNLGSFSELSSVSSVSSGTEGASSLEKKEVP 780
QY 947 GVDFTSQFVRNLGLEHMLDIFEREQITLDVLEMGHKEKELKEIGINAYGHRHKLKIGVER 1006
Db 781 GVDFTSQFVRNLGLEHMLDIFEREQITLDVLEMGHKEKELKEIGINAYGHRHKLKIGVER 840
QY 1007 LISGQOGLNPLYLNTSGSGTILIDLSPPDKFQSVSEEMQSVREHRRDGGHAGGIFNRY 1066
Db 841 LISGQOGLNPLYLNTSGSGTILIDLSPPDKFQSVSEEMQSVREHRRDGGHAGGIFNRY 900
QY 1067 NILKIOKVCNKKLWERYTHRRKEVSEENHNANERMLFHGSPFVNATIHKGDFERHAYIG 1126
Db 901 NILKIOKVCNKKLWERYTHRRKEVSEENHNANERMLFHGSPFVNATIHKGDFERHAYIG 960
QY 1127 GMFGAGIYFAENSCKSNQYVYIGGGTGCPCVHKDRSCYICHRQLLCFVTLGKSFQFSA 1186
Db 961 GMFGAGIYFAENSCKSNQYVYIGGGTGCPCVHKDRSCYICHRQLLCFVTLGKSFQFSA 1020
QY 1187 MKMAHSPGHHSTGPRVSNGLALAEYVIYRGQAPPEYLITVQIMRPEGWDG 1240
Db 1021 MKMAHSPGHHSTGPRVSNGLALAEYVIYRGQAPPEYLITVQIMRPEGWDG 1074

RESULT 2

US-09-841-835-2
; Sequence 2, Application US/09841835
; Patent No. US20020076795A1

GENERAL INFORMATION:

; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRP1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack

; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:

CLASSIFICATION:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521

INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1327 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOPHETICAL: NO

US-09-841-835-2

Query Match 79.0%; Score 5103.5; DB 10; Length 1327;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 963; Conservative 110; Mismatches 128; Indels 15; Gaps 3;

QY 22 AAGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRCA 81
Db 112 SAAGVAPNPAGSGNNSPSSSPTSSSSSSPSPG-----SSLAESPAAAGVSSSTAPL 165
QY 82 GGAACAASAAAEAVEPAARELFEACRNGDVERVKRIVTPEKVNSTRDTAGKRKSTPLHPAAG 141
Db 166 GPGAAAGTGVPAVSGALRELLACRNGDVSRYKRLVDAANVNAKMDAGRKSSPLHPAAG 225
QY 142 FGRKDVVEYLLQNGANVOARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNYTP 201
Db 226 FGRKDVVEYLLQNGANVHARDGGLIPLHNACSFHAEVNVNLLRCCQADPNARDNNYTP 285
QY 202 LHEAAITKGIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKKDELLESARS 261
Db 286 LHEAAITKGIDVCIVLLQHGADPNIRNTDGTALDLADPSAKAVLTGEYKKDELLESARS 345
QY 262 GNEKEMWALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLV 321
Db 346 GNEKEMWALLTPLNVNCHASDGRKSTPLHLAAGYNRVIVOLLQHGADVHAKDKGLV 405
QY 322 LHNACSYGHEVTELLVKGACVNAAMDWQFTPLHAAASKNRVEVCSSLISYGADPTLLN 381
Db 406 LHNACSYGHEVTELLVKGACVNAAMDWQFTPLHAAASKNRVEVCSSLISYGADPTLLN 465
QY 382 CHNKAIDLAAPTQPKERLAYEFGKHSLLQAAAREADVTRIKKHLSELMVNFKHPQTHETA 441
Db 466 CHKSAVDNAPTPELRERLTTEFGKHSLLQAAAREADLAKVKKTLALEIINFKQPQSHETA 525
QY 442 LHCAASAPYKPKQICELLRLKCANINEKTEKFTPLPLHVASEKAHNDVVVVVVKHAKVN 501
Db 526 LHCAVASLHPKQVQVTELLRLKCANINEKTEKFTPLPLHVASEKAHNDVVVVVVKHAKVN 585
QY 502 ALDNLGOTSLHRAAYCGHLQTCRLLISYGDPNIIISLQGTALQMGNNVQOQLQEGISL 561
Db 586 ALDNLGOTSLHRAAYCGHLQTCRLLISYGDPSIIISLQGTALQMGNEAVQOILSESTPI 645

QY 562 GNSEADROLLEAAKAGDVETVKKLTQVSQVNCRDIEGRQSTPLHFAAGYNNRVSVVEYLQ 621
Db 646 ITSDVDYRLLEAKAGDLETVKQLCQSSQVNCVNCRLDLEGRHSTPLHFAAGYNNRVSVVEYLH 705
QY 622 HCADVHAKDKGLVPLHNAACSGHYEVAELLYKHGAVNVADLWKFTPLHFAAAGKGYEI 681
Db 706 HCADVHAKDKGLVPLHNAACSGHYEVAELLYKHGAVNVADLWKFTPLHFAAAGKGYEI 765
QY 682 CKLLQHGADPTKKNRDGNTPLDLVKDGTDTQDILLRGGDAALLDAACKGCLARVKKLSPP 741
Db 766 CKLLKHGADPTKKNRDGNTPLDLVKEGDTQDILLRGGDAALLDAACKGCLARVKKLTCP 825
QY 742 DNVNCRDTQGRHSTPLHNAAGYNNLEVAEYLQHGADVNAQDKGLIPLHNAASGYCHVDV 801
Db 826 ENINCRDTQGRNSTPLHNAAGYNNLEVAEYLLEHGADVNAQDKGLIPLHNAASGYCHVDI 885
QY 802 AALLIKYNACVNATDKWFTPLHFAAQKRTOLCALLLAHAGADPTLKNOEGOTPLDLVSA 861
Db 886 AALLIKYNACVNATDKWFTPLHFAAQKRTOLCALLLAHAGADPTLKNOEGOTPLDLVSA 945
QY 862 DDVSALLTAAMPSPALPCYKQVNLGVKSPGATADALSSGSPSSLSAASSLNLGSG 921
Db 946 DDIRALLIDAMPPEALPTCFKPKQAT-----VVSASLISPASTPSCLSAASSIDNLGTP 998
QY 922 FSLSVSSVSSCTEGASSLEKK--EVPQVDFSIOTFVRNLGLEHMDIFEREQITLDVLV 979
Db 999 LAELAVGASNAAGDGAAGTERKEGEVAGLDNMISOFKSLGLEHLRDIPEFETQITLDVLA 1058
QY 980 EMGHKEKEIGINAYGHRKHLKFGVERLTSGOGLNPYLTNTSGTILDLSPDKKEF 1039
Db 1059 DMGHKEKEIGINAYGHRKHLKFGVERLTSGOGLNPYLTNTSGTILDLSPDKKEF 1118
QY 1040 OSVEEMQSTVREHRDGGHAGGIFNRYNLIKQVCKNKKLWERYTHRRKEVSEENHNHAN 1099
Db 1119 OSVEEMQSTVREHRDGGHAGGIFNRYNLIKQVCKNKKLWERYTHRRKEVSEENHNHAN 1178
QY 1100 ERMFLPHGSPFVNAIITHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYIGGGTGCPCVHK 1159
Db 1179 ERMFLPHGSPFVNAIITHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYIGGGTGCPCVHK 1238
QY 1160 DRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLAELAEYVYIRGE 1219
Db 1239 DRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLAELAEYVYIRGE 1298
QY 1220 QAYPEYLITYQIMRPE 1235
Db 1299 QAYPEYLITYQIMRPE 1314

RESULT 3
US-09-964-899-41
; Sequence 41, Application US/09964899
; Patent No. US2002017446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-41

Query Match

60.9%; Score 3938; DB 9; Length 802;

Best Local Similarity 67.2%; Pred. No. 1.9e-270;
Matches 787; Conservative 4; Mismatches 5; Indels 376; Gaps 5;
QY 75 MSGRRCAGGAGAACASAAAFAVEPAARELFEACRNGDVERVKRLVTPPEKVNSTRDTAGRKST 134
Db 1 MSGRRCAGGAGAACASAAAFAVEPAARELFEACRNGDVERVKRLVTPPEKVNSTRDTAGRKST 60
QY 135 PLHFAAGFRKDVVEYLLONGANVOARDGGLIPLHNAACSGHYEVAELLYKHGADPNAR 194
Db 61 PLHFAAGFRKDVVEYLLONGANVOARDGGLIPLHNAACSGHYEVAELLYKHGADPNAR 120
QY 195 DNWNTPLHEAATKGDVCIIVLLQHGABPTIRNTDGRALTDLADPSAKAVLTGE----- 249
Db 121 DNWNTPLHEAATKGDVCIIVLLQHGABPTIRNTDGRALTDLADPSAKAVLTGKSVYS 180
QY 250 YKKDELLES-ARSGNEEKMMALLTPLNVNCHASDGRKSTPLHNAAGYNNRVKIVOLLQHG 308
Db 181 YSRKPKVKNLARSNEEKMMALLTPLNVNCHASDGR----- 216
QY 309 ADVHAKDKGLVPLHNAACSGHYEVEVELLVKHGACVNAAMDLMQFTPLHFAASKNRVEVCS 368
Db 217 -----KHGACVNAAMDLMQFTPLHFAASKNRVEVCS 246
QY 369 LLSYGADPTLLNCHNKSALDAPTPQLKERLAYEFKHSLLQAAAREADVTRIKKHLISLE 428
Db 247 LLSYGADPTLLNCHNKSALDAPTPQLKERLAYEFKHSLLQAAAREADVTRIKKHLISLE 306
QY 429 MYNFKHPQTHETALHCAASAPYKPKKOICEALLRKGANINEKTEKFTPLHVAASEKAHND 488
Db 307 MYNFKHPQTHETAL----- 320
QY 489 VVEVVVVKHEAKNALDNLGQTSILHRAAYCGHLQTCRLLLSYGDPNIIISLOGFTALQMG 548
Db 321 ----- 320
QY 549 ENVQQLLOBEGISLGNSEADROLLEAAKAGDVETVKKLTQVSQVNCRDIEGRQSTPLHFAA 608
Db 321 -----KLCTVQSVNCRDIEGRQSTPLHFAA 345
QY 609 GYNNRVSVVEYLLQHGADVHAKDKGLVPLHNAACSGHYEVAELLYKHGAVNVADLWKFT 668
Db 346 GYNNRVSVVEYLLQHGADVHAKDKGLVPLHNAACSGHYEVAELLYKHGAVNVADLWKFT 405
QY 669 PLHEAAAAGKYETICLLQHGADPTKKNRDGNTPLDLVKDGTDTQDILLRGGDAALLDAK 728
Db 406 PLHEAAAAGKYETICLLQHGADPTKKNRDGNTPLDLVKDGTDTQDILLRGGDAALLDAK 465
QY 729 KGCLARVKKLSSPDNVNCRDTQGRHSTPLHNAAGYNNLEVAEYLQHGADVNAQDKGGLI 788
Db 466 KGCLARVKKLSSPDNVNCRDTQGRHSTPLHNAAGYNNLEVAEYLQHGADVNAQDKGGLI 525
QY 789 PLHNAASYCHVDVVAALLIKYNACVNATDKWFTPLHFAAQKRTOLCALLLAHAGADPTLK 848
Db 526 PLHNAASYG----- 534
QY 849 NOEGOTPLDLVSADDDVALLTAAMPSPALPCYKQVNLGVKSPGATADALSSGSPSS 908
Db 535 ----- 534
QY 909 LSAASSLNLGSGFSSELSSVSSSCTEGASSLEKKEVPGVDFSIOTFVRNLGLEHMDIF 968
Db 535 ----- 534
QY 969 EREQITLDVLMVGMHKEKEIGINAYGHRKHLKFGVERLTSGOGLNPYLTNTSGGTI 1028
Db 535 -----ITLDVLMVGMHKEKEIGINAYGHRKHLKFGVERLTSGOGLNPYLTNTSGGTI 590
QY 1029 LIDLSPDKKEFOSVEEMQSTVREHRDGGHAGGIFNRYNLIKQVCKNKKLWERYTHRRK 1088
Db 591 LIDLSPDKKEFOSVEEMQSTVREHRDGGHAGGIFNRYNLIKQVCKNKKLWERYTHRRK 650
QY 1089 EYSEENHNHANERMLFHGSPFVNAIITHKGFDERHAYIGMFGAGIYFAENSSKSNQYVY 1148
Db 1089 EYSEENHNHANERMLFHGSPFVNAIITHKGFDERHAYIGMFGAGIYFAENSSKSNQYVY 1148

Db 651 EVSEENHNANERMLPHGSPFVNATIIHKGFDERHAYIGMFGAGIYFAENSSKSNQVYVG 710
QY 1149 IGGTCCPVHKDRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPPGHSHVTVGRPSVNGL 1208
Db 711 IGGTCCPVHKDRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPPGHSHVTVGRPSVNGL 770
QY 1209 ALAEYVYIRGEQAYPEYLITYQIMRPEGMVDG 1240
Db 771 ALAEYVYIRGEQAYPEYLITYQIMRPEGMVDG 802

RESULT 4
US-09-841-835-10
; Sequence 10, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 949 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-841-835-10

Query Match 51.2%; Score 3307; DB 10; Length 949;
Best Local Similarity 76.2%; Pred. No. 1e-225;
Matches 638; Conservative 72; Mismatches 119; Indels 8; Gaps 2;

QY 22 AAGTADPVTAGSQARALSSPGGLALLAGPGLLLRLALLLAVAAARIMSGRCA 81
Db 112 SAAGVAPNAGSGNNSSPSSTSSSSPSPG-----SSLAESPEAAGVSTAPL 165
QY 82 GGAACASAAAEAVEPAARELFACRNGDVERVKRLVTPKVNSRDTAGRKSTPLHFAAG 141
Db 166 GPGAAGGTGVAPVAGRLRELACRNGDVRVKRLVDAANNAKMDAGRKSSPLHFAAG 225
QY 142 FGRKDVVYLLQNGANVQARDGGLIPLHNACSGFHAENVNLLRHGADPNARDNNWYTP 201
Db 226 FGRKDVVYLLQNGANVHARDGGLIPLHNACSGFHAENVNLLRCQADPNARDNNWYTP 285
QY 202 LHEAAIKGKIDVICIVLQHGADPNIIRNTDGTALDADPSAKAVLTGEYKKDELLESARS 261

Db 286 LHEAAIKGKIDVICIVLQHGADPNIIRNTDGTALDADPSAKAVLTGEYKKDELLESARS 345
QY 262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLVP 321
Db 346 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLVP 405
QY 322 LHNACSYGHYEVTELLVKHGACVNMALMQFTPLHFAASKNRVEVCSLLLSYGADPTLN 381
Db 406 LHNACSYGHYEVTELLVKHGACVNMALMQFTPLHFAASKNRVEVCSLLLSHGADPTLN 465
QY 382 CHNKSALDAPTPOLKERTAYEFKSHLSLOAAREADVTRIKKHLSEMVNFKHPQTHETA 441
Db 466 CHKSASVDMAPTDELRLTYEFKSHLSLOAAREADLAKYKKTALETIIRFKQPQSHETA 525
QY 442 LHCAASPYPKKOICELLRLKGANINEKTEFLTPLVASEKAHNDVVEVYVYKHAQVN 501
Db 526 LHCASVSLHPKRVQVTELLRLKGANYNKNDPMTPLHVAERAHNDVMEVLKHGAKMN 585
QY 502 ALDNLGQTSIHLAAYCGHLQTCRLLLSYGCDPNTIISLQGTALOMGNENYQQLQEGISL 561
Db 586 ALDTLQGTALHRAALAGHLQTCRLLLSYGSDPSIISLQGTAAQMGNEAVQQLISESTPI 645
QY 562 GNSEADPQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVWEYLLQ 621
Db 646 RTSVDVYRLLEASKAGDLETVKLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVWEYLLH 705
QY 622 HGADVHAKDKGLVPPLHNACSYGHYEAELLVKHGVVNVADLWKFTPLHFAAAGKYEI 681
Db 706 HGADVHAKDKGLVPPLHNACSYGHYEAELLVHGHASVNVADLWKFTPLHFAAAGKYEI 765
QY 682 CKLLQHGADPTKKNRDGNTPLDLVKDGTDIQDLLRGDAALLDAAKKGCLARVKLSSP 741
Db 766 CKLLKHGADPTKKNRDGNTPLDLVKEGDTIQDLLKGDAAALLDAAKKGCLARVKLCTP 825
QY 742 DNVNCRDTQGRHSTPLHLAAGYNNEVLAEYLLQHGADVNAQDKGGLIPLHNAASYGH--V 799
Db 826 ENINCRDTQGRNSTPLHLAAGYNNEVLAEYLLHGHADVNAQDKGGLIPLHNAASYGCLA 885
QY 800 DVAALLIKYNACVNAETDKWAFTPLEHAAKQRTQCALLLAHAGADPTLKNQEGQTP 856
Db 886 RVOKLCTPENINCRDTQGRNSTPLHLAAGYNNEVLAEYLLHGHADVNAQDKGGLIPL 942

RESULT 5
US-09-841-835-8
; Sequence 8, Application us/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:

Db 78 LHLASKGHEVVEVSELLQREANVDAAATKGN TALHIASLAGAEVVKVLYVTNGANVNAQS 137
QY 196 NWNVTPLHEAAIKIDVCIVLLQHGAEPTIRNTDGRDTALDLPADPSAKAVLTGEYKKDEL 255
Db 138 QNGFTPLYMAXQENHLEWVFLDNGASQXLXTEDGTPLAVA----- 180
QY 256 LESARSNEEKMMALLTPLNVNCHASGRKSTP-LHLAAGVNRVKIVOLLQHGADVHAK 314
Db 181 ---LQOQHDQVVSLLLE-----NDTKGKVRPLPALHTAARKDDTKAAALLIQDNNDNADVE 231
QY 315 DKGLDPLVHNACSYGHEVTELLVKKHCACYNAMDLPQFTPLHEAASKNRVEVCSLLLSYG 374
Db 232 SKSGFTPLHTAAHGNINATLLNRAAVDFTARNIDTPLHVASKRGANMMVKLLLDGR 291
QY 375 ADPTLLNCHNKSAIDLAPTQPKERLAYEFKGHSLLOAAAREADVTRIKKHLSLEWVNFKH 434
Db 292 A-----KIDAKTRGCL----- 302
QY 435 POTHETALHCAASPYPKRQOICELLRKGANINEKTEFTPLHVASEKAHNDVVEVV 494
Db 303 ---TPHCGGARGH---EQVVEMLLDRAAPILSKTKNGLSPLHMATQGDHLNCVQLL 354
QY 495 KHEAKVNALDNGOTSLHRAAYCGHLOTCRLL 527
Db 355 OHNVPPVDVNTDNYLTALHVAHCHGYKAVKVL 387

RESULT 8

US-10-012-896-378

; Sequence 378, Application US/10012896

; Publication No. US20020183251A1

; GENERAL INFORMATION:

; APPLICANT: Xu, JiangChun

; APPLICANT: Dillion, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Meagher, Madeleine Joy

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.427C27

; CURRENT APPLICATION NUMBER: US/10/012.896

; CURRENT FILING DATE: 2001-12-10

; NUMBER OF SEQ ID NOS: 1011

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 378

; LENGTH: 1719

; TYPE: PR

; ORGANISM: Homo sapiens

US-10-012-896-378

Query Match

Best Local Similarity 6.2%; Score 398.5; DB 9; Length 1719;

Matches 237; Conservative 137; Mismatches 411; Indels 327; Gaps 34;

QY 80 CAGGAACASAAAEAVEPAAREL-----FEACRNGDVERV-----KRLVTP 120
Db 72 CRGSGSKNVGASGDHDDSAMKTLRNKMGKWCCHCPCPCRGSGSKVGAWDYDSDAFMEP 131
QY 121 EKVNSRDTAGKSTPLHFAAGFG---RKDVVEYLQNGANYOARDGGLIPLHNAACSGH 177
Db 132 R-----YHVRGEDLDKLHRAAWGKVPKRDLI--VMLRDTDVNKKDKRKTALHLASANG 185
QY 178 AEVYNLLLRHGCADPNARDNWNVTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRDTALD 237
Db 186 SEVVKLLLDRCQLNVLDNKKRTALIKAVQCEDEBCALMLLEHGDPNIPDEYGNWTTLHY 245
QY 238 ADPSAKAVLTGEYKKDELLESARSNEEKMMALLTPLNVNCHASGRKSTPLHLAAGVNR 297
Db 246 A-----TYNEDKLMA----- 255
QY 298 VKTVOLLQHGADVHAKDKGLVPLHNACSYGHEVTELLVKKHCACYNAMDLPQFTPLHE 357
Db 256 ---KALLYGADIESKNKHGLTPLLGVHEQKQOVVKFLIKKANLANALDRYGR TALIL 311
QY 358 AASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPT--QLKERLAYEFKGHSLLOAAARE 415
Db 312 AVCCGSASIVSLLLEQNTDVSSQDLSGQTAREYAVSSHVHCICQLLSYKEKOMLUKISE 371
QY 416 ---ADVTRIKKHLSLEMVNFKHPQTH---ETALHCAAS---PYPKRQOICELLRLRKA 465
Db 372 NSNPENVSRT-------NKPRTHMVVEVDSMPAASSVKKPGPLRSKMGKWCRCRCP 421
QY 466 NINEKTEFTPLHVASEKAHNDVVEVVVHKEAKVNALDNGOTSLHRAAYCGHLOTCRL 525
Db 422 CCRESGKS---NVGTSGDHDD-----SAMKTLRSKMGK-----WCRHCFPC-- 459
QY 526 LLSYCCDPNIISLOGFTALQMNENVOQLLOGLISLGNSEADROLLEAAKAGDVETVKKL 585
Db 460 ---CRGS-----GKSNV-----GASGDHDDSAMKTLR-----NKMKGW 489
QY 586 CTQVSVNCR-----DIEGRSTPLHFAAGVNRYSVVEYLLQ- 621
Db 490 CCHCFPCRGSGSKVGAWDYDSDAFMEPRYHVRGEDLDKLHRAAWGKVPKRDLIYVL 549
QY 622 HGADVHAKDKGGLVPLHNACSYGHEVVAELLVKHGAVNVADLWKFTPLHRAAAKGYEI 681
Db 550 RDTDVNKKDKOKRTALHLASANGSEVVKLLDRRCQLNVLDNKKRTALIKAVQCEDEC 609
QY 682 CKLLLOHCADPTKKNRGNTPLDLVKDGTDTIDQLLRGDAALLDAKKGCLARVKKLSPP 741
Db 610 ALMLLEHGTDPNIPDEYGN----- 628
QY 742 DNVNCRDTQGRHSTPLHLAAGVNNLEVAEYLLQHGADVNAODKGLIPLHNAASYGHVDV 801
Db 629 -----TTLHYAIYNEDKLMAKALLYGADIESKNKHGLTPLLGVHEQKQOV 675
QY 802 AALLIKYNACVNATDKWAFTPLHEAAQKRTOLCALLAHAGADPTLKNQEGOTPLDLVSA 861
Db 676 VKFLIKKANLANALDRYGR TALILAVCCGSASIVSLLLEQNTDVSSQDLSGQTAREYAVS 735
QY 862 DDVSALLTAAMPSPNALPSCYKPOVLNGVSPGATADALSSGSPSSLSAASLDNLSGS 921
Db 736 SHHVI-----COLJSDYKEKOMLKISSE-----NSNPEQDLKLTSEESORFVGS 781
QY 922 FSELSSVVSSSGTEGASLEKKEVPGVDFSTITQFVRNLGLEHLMDFEREQITLVDVEM 981
Db 782 -----ENSQPEKMSQPEINKD-----GPRE-----VE- 804
QY 982 GHKELKEIGINAYGHRHKLIGVERLISGQOQLNPYLTNTSGSTILIDJSPDDKEFQS 1041
Db 805 ---EEMKKHESNNVGLLENLTNGV-TAGNCDNGLIPORKSRT-----PENQOFPD 850
QY 1042 VEEEMQSTVREHRRDGGHAGGIFNRYNILKIOKVCNK-----KLWERYTHRRKEVSEE 1093
Db 851 NESEYHRICE-----LVSDYKEKOMPKYSSSENPEQDLKLTSEESORLEGSEN 901
QY 1094 NHHN-----ANERMLFHGSGPFVNAIIHKGFDE 1120

Db 902 GOPELENFMAIEEMKKHGS-----THVGFPPE 927

RESULT 9
US-09-895-793-378
; Sequence 378, Application US/09895793
; Publication NO. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 378
; LENGTH: 1719
; TYPE: PRF
; ORGANISM: Homo sapien
US-09-895-793-378

Query Match 6.2%; Score 398.5; DB 9; Length 1719;
Best Local Similarity 21.3%; Pred. No. 1.4e-19;
Matches 237; Conservative 137; Mismatches 411; Indels 327; Gaps 34;
QY 80 CAGGGACASAAAEVPEAREL-----FEACRNGDVERV-----KRLVTP 120
Db 72 CRGSGSNVGVAGDHDSDAMKTLRNKMGKWCCHCFPCCRGSGKSKYGVANGDYDDSAFMEP 131
QY 121 EKVNSRDTAGKSTPLHFAAGFG---RKDVVEYLLQNGANVQARDGGLIPLHNACSFHG 177
Db 132 R----YHVRGEDLDKLHRAAWGKVPKDLI--VMLRDTDVNKKQKQKRTALHLASANGN 185
QY 178 AEVNNLLRHGADPNARNWNYPPLHEAAIKGKIDVICVILLQHGAEPTIRNTDGRATDLD 237
Db 186 SEVVKLLDRRCOLNLDNKKRTALIKAVQCQDECALMLEHGTDPNIPDEYGNTHLY 245
QY 238 ADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDRKSTPLHLAAGYNR 297
Db 246 A-----IYNEDKMLA----- 255
QY 298 VKTVOLLQHGADVHAKDGLVPLHNACSYGHYEYTELLVKGACVNMALWQFTPLHE 357
Db 256 ----KALLYGADIESKNHGLTPLLGVHEQKQVVKELIKKKANLNALDRYGRATALIL 311
QY 358 ASKNRVCSLLLSYGADPTLLNCHINKSAIDLAPTP--OLKRLAYEYFKGSHSLQAAARE 415
Db 312 AVCCGSASIVSKLLLEQNIIDVSSODLSGOTAREYAVSSHVHVICQLLSDYKEKOMLKISSE 371
QY 416 ----ADVTRIKKHLSEMYNFKHPQTH---ETALHCAAS---PYPRKQICELLRLKGA 465

Db 372 NSNPENVSRT-----NKPRTHMVVEVDSMAASVKKPFGGLRSMKMGWCRCFP 421
QY 466 NINEKTKFELTPLHVASAKAHNDVVEVVKHEAKVNALDNLGOTSLHRAAYCCHLOTCLR 525
Db 422 CCRSECKS-----NVGTSGDHDD-----SAMKTLRSKMKG-----WCRHCFPC-- 459
QY 526 LLSYGCDPNIIISLQGFALQMGNNENVOQLQEGISLGNSEADROLLEAAKAGDVETVKKL 585
Db 460 -----CRGS-----GKSNV-----CASGDHDDSAMKTLR-----NKMKGW 489
QY 586 CIVQSVNCR-----DIEGRQSTPLHFAAGYNRVSVVEYLLQ-- 621
Db 490 CCHCFPCCRGSGKSKYGVANGDYDDSAFMEPRYHVRGEDLDKLHRAAWGKVPKDLIVML 549
QY 622 HGADVHAKDKGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKETPLHEAAAKGKYEI 681
Db 550 RDTDVNKKQKRTALHLASANGSEVVKLLDRRCOLNLDNKKRTALIKAVQCQDEDEC 609
QY 682 CKLLQHGADPTKKNRDGNTPLDLVKDGTDIQDILLRGDAALLDAAKKGCLARVKKLSSP 741
Db 610 ALMLEHGTDPNIPDEYN----- 628
QY 742 DNVNCRDTCGRHSTPLHLAAGYNNLEVAEYLLQHGADVNNADKGGGLIPLHNAASYGHVDV 801
Db 629 -----TTLHYAIYNEDKLMKALLYGADIESKNKHGLTPLLGLVHEQKQV 675
QY 802 AALLIKYNACVNATDKWATPLHEAAQKGRTOCALLLAHGADPTLKNQBOGTPDLVSA 861
Db 676 VFELIKKANLNALDRYGRATALILAVCCSASIVSLLLEQNIIDVSSODLSGOTAREYAVS 735
QY 862 DVVSALLTAAMPPSALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASLDNLGSG 921
Db 736 SHHVI-----CQLLSDYKEKOMLKISSE-----NSNPEQDLKLTSEESQRFKGS 781
QY 922 FSELSSVSSSGTEGASLSLEKKEVPGVDFPSITOFVRNLGLEHLMDFEREQITLDVLVEM 981
Db 782 -----ENSQPEKMSQEPENKD-----GDRE-----VE- 804
QY 982 GHKELKEIGINAYGHRHKLTKGVERLISGQGLNPVLTNTSGTILIDLSPDDKEFQS 1041
Db 805 --EEMKKHESNNVGLLENLTNGV-TAGNDGNGLIPOKRSPT-----PENQOFPD 850
QY 1042 VEEMOSTVREHRDGHGAGIFNRYNILLIKQKVCNK-----KLWERYTHRRKEVSEE 1093
Db 851 NSEEEYHRICE-----LVSDYKEKOMPKYSSSENSNPQDLKLTSEESQORLEGSEN 901
QY 1094 NNNH-----ANERMLFHSFPFVNAIHKGFDE 1120
Db 902 GOPELENFMAIEEMKKHGS-----THVGFPPE 927

RESULT 10
US-09-895-814-378
; Sequence 378, Application US/09895814
; Publication NO. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.

```

; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 378
; LENGTH: 1719
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-895-814-378

Query Match
Best Local Similarity 21.3%; Score 398.5; DB 9; Length 1719;
Matches 237; Conservative 137; Mismatches 411; Indels 327; Gaps 34;

Qy 80 CAGGGAACASAAAEVPAAREL-----FEACRNGDVERV-----KRLVTP 120
Db 72 CRGSGKSNVAGSDHDDDSAMKTLRNKMKWCCHCFCCRGSGKSKVGAWGDYDDSAFMEP 131

Qy 121 EKVNSRDTAGRKSTPLHFAAGFG---RKDVVEYLLQNGANVOARDGGGLIPLHNACSF 177
Db 132 R-----YHVRGDDLDKLRHAAWGWKVPKDLI--VMLRDTVNNKDKOKRKTALHLASANGN 185

Qy 178 AEVYNLLRHGADPNARDNNMNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRALDL 237
Db 186 SEVVKLLDRRCQLNLDNKKRTALIKAVQCOEDCALMLLEHGTDPNIPDEYGN'TLHY 245

Qy 238 ADPSAKAVLTGEYKKDELLESARGSENEKMWALLTPLNVNCHASDGRKSTPLHLAAGNVR 297
Db 246 A-----TYNEDKLMA----- 255

Qy 298 VKIVQLLQHGADVHAKDGLVPLHNACSYGHYEVTLLVKHCACVNAMDLWOFTPLHE 357
Db 256 ----KALLYGADIESKNKCLTPLLGVHFOKQOVVKFLIKKANLALNDRYGRTALIL 311

Qy 358 AASKNRVEVCSLLSYGADPTLLNCHNKSADLAPT--QLKERLAYEFKHSLLQAARE 415
Db 312 AVCCGSASIVSLLLEQNIIDVSSQDLSGQTAREYAVSSHVVHVICQLLSDYKEKOMLKISSE 371

Qy 416 ----ADVTRIKKLSLEWVNFKHPQTH---ETALHCAAS---PYPRKQKICELLLRKGA 465
Db 372 NSNPENVSRT-----NKPRTHMVVEVDSMPAASVYKPKFGLRSKMGKWCRCRFP 421

Qy 466 NINEKTRFTPLHVAKEAHNDVVEVVKHAEKVNALDNLGOTSLHRAACVGHLOTRCL 525
Db 422 CCRESGKS-----NVGTSGDHD-----SANKTLRSKMG-----WCRHCFPC-- 459

Qy 526 LLSYGCDPNISLQGTALQMNENVOQLQEGISLGNSEADROLLEAAKAGDVETVKKL 585
Db 460 ----CRGS-----GKSNV-----GASGDHDDDSAMKTLR-----NKGWK 489

Qy 586 CTQVSVNCR-----DIEGRQSTPLHFAAGYNRVSVVYLLQ-- 621
Db 490 CCHCFCCCRGSGKSKVGAWGDYDDSAFMEPRYHVRGDDLDKLRHAAWGWKVPKRDILVML 549

Qy 622 HGADVHAKDGGGLVPLHNACSYGHYEAVELLVKGAVVNVADLWKFTPLHEAAAKGYEI 681
Db 550 RDTVNNKDKOKRKTALHLASANGSEVVKLLDRRCQLNLDNKKRTALIKAVQCOBDEC 609

Qy 682 CKLLQHGADPTKKNRDNPTPLDIAVKDGDTHDIQDLRGDAALLDAARKGGLARVKKLSSP 741
Db 610 ALMLEHCTDPNIPDEYGN----- 628

Qy 742 DNVNCRDTQGRHSTPLHLAAGYNMLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801
Db 132 R-----YHVRGDDLDKLRHAAWGWKVPKDLI--VMLRDTVNNKDKOKRKTALHLASANGN 185

```

Qy	178	AEVYNLLLRHGADPNARDWNVYTPLEHAAIKGIDVCIVLIQHGAEPTRINTDGR	TDL	237
		:	:	
Db	186	SEVVKLLLRROCLNDLNKRTALIKAYOCQEDCALMLLEHGTGDPNIPDEYNTIHY		245
Qy	238	ADPSAKAVLTGEYKDBELLESARSGNEEKMAILLPLNVNCHASDGRKSTPLHLAAGYNR		297
Db	246	A-----IYNEDKMA-----		255
Qy	298	VKIVQLLQHGADVHAKDGDLVPLHNACSYGHYEVELLVKHGACVNMADLWQFTPLHE		357
Db	256	---KALLYGADIESKNHGLTPLLGVHEKQOQVWFLLIKKANLANLADRYGRTALIL		311
Qy	358	AASKNRVECSLLSYCADPTLLNCHNSAIDLATP--OLKERLAYEFKHSLLQARE		415
Db	312	AVCGSASIVSLLLEONIDVSSQDLSGQTARYAVSSHHHVTCQLSDYKEKQMLKTSSE		371
Qy	416	---ADVTRIKKLSLEWYFNKHQPTH---ETALHCAAS---PYPRKPKICEJLLLRKA		465
Db	372	NSNPENVSRT-----NKPRTHMVVEDSMPAASSVKKPGFLRSKMGKWCRCFP		421
Qy	466	NINEKTEFTPLHVASEKAHNDVVEVVKHEAKYNALDNLGQTSLHRAAYCGHLQTCRL		525
Db	422	CORESGKS-----NVGTSGDHD-----SAMKTLRSKMKG-----WCRHCFPC--		459
Qy	526	LLSYGCDPNIIISLQGTALQMGNENVQQLQEGISLGNSEADROLLEAAKAGDVETVKL		585
Db	460	----CHGS-----GKSNV-----GASGDHDSAMKTLR-----NKMGK		489
Qy	586	CTVOSVNCR-----DIEGROSTPLHFAAGYNRVSVVEYLLQ--		621
Db	490	CHCFPCCRSGSKSGVGWGDYDSDAFMEPRYHVRGEDLDKLHRAAWMGKVPKRDLI		549
Qy	622	HGADVHAKDGGGLVPLHNACSYGHVEAELLVKHGAVNVADLWKFTPLHEAAAKGYEI		681
Db	550	RDTVYKKDKKRTALHLASANGSEVYKLLDRRCQNLVLDNKKRTALIIRAVOCQEDC		609
Qy	682	CKLLQHGADPTKKNRDGNTPDLVKDGPDTDODLLRGDAALLDAAKGCLARYKKLSSP		741
Db	610	ALMLLEHGTDPNIPDEYN-----		628
Qy	742	DNVNCRDQTGRHSTPLHLAAGYNLEVAEYLLQHGADYNQODKGLIPLHNAASYGHVDV		801
Db	629	-----TTHYAIYNEDKLMAKALLYGADIESKNKHGLTPLLGVHEQKQOV		675
Qy	802	AALLIKYNACVNATDKWAFPTLHEAAQKQRTOLCALLHAGADPTLKNQEGTPLDLVSA		861
Db	676	VKFLIKKANLANLADRYGRTALILAVCCGSASIVSLLLEQNIDVSSQDLSGOTAREYAVS		735
Qy	862	DDVSALLTAAMPSPALPSYCKQVNLGVRSQATADALSSGSPSPSSLSAASSIDNLISGS		921
Db	736	SHHHVI-----COLSDYKKEQMLKTSSE-----NSNPEQDKLITSEESQRFKGS		781
Qy	922	FSELSSVSSSGTEGASSLEKKEVPGVDSITQFVRNLGLEHLMDFIREQITLIDLIVEM		981
Db	782	-----ENSOPEKMSQEPINK-----GDRE-----VE--		804
Qy	982	GHKELKEIGINAYCHRRHLKIGVERLISQOQGLNPLYLINTSGSTILIDLSPDKCFQS		1041
Db	805	--EEMKKHESNNYGLLENLTNGV-TAGNDGNGLIIPQRKST-----PENQOFPD		850
Qy	1042	VEEMQSTVREHRDGGHAGGIFNRVYMLKIQKVCNK-----KLWERYTHRKEVSEE		1093
Db	851	NESEYHRICE-----LVS DYKEKOMPYSSENSNPQDCLKLTSEESQRLGSEN		901
Qy	1094	NNHH-----ANERMLPHGSPFVNAIHKGFDE		1120
Db	902	GQPELENFMAIEMKKHGS-----THVGFE		927

RESULT 12
US-09-780-669-378
: Sequence 378, Application US/09780669

```

; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 378
; LENGTH: 1719
; TYPE: prt
; ORGANISM: Homo sapien
; US-09-780-669-378

```

Query Match	6.2%	Score 398.5;	DB 10;	Length 1719;
Best Local Similarity	21.3%;	Pred. No. 1.4e-19;		
Matches 237;	Conservative 137;	Mismatches 411;	Indels 327;	Gaps
QY	80	CAGGGAACASAAAEAVPAAREL-----FEACRNGDVERV-----KRLVTP 120		
Db	72	CRSGSKNVGASGDHDSAMKTLRNKMGKWCCHCFCCRGSGSKGVANGWDYDSAFMEP 131		
QY	121	EKVSNRDTAGRKSTPLHFAAGF---RKQVYEXLLONGANVQARDGGGLPLHNACSFHG 177		
Db	132	R-----YHVRGEDLDKLHRAAWMGVPRKDLI---VMLRDTDVNKKDKQKRTALHLASANG 185		
QY	178	AEVYNLLLRHCADPNARDNNYTPLEHAAIKGKLDVCIVLLQHGAEPTIRNTDGR TALDL 237		
Db	186	SEVVKLLLRRCQINVLNDNKKRTALIKAVOCDECALMLLEHGTOPNIPDEYGN TTHY 245		
QY	238	ADPSAKAVLTGEYKKDELLESARGNEEKMAALLTPLNVNCHASDGRKSTPLHLAAGYNR 297		
Db	246	A-----LYNEDKLMA----- 255		
QY	298	VKIVOLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKKHGACVNAMDLMQFTPLHE 357		
Db	256	-----KALLLYGADTESKNKHGLTPLLLGVHEQOQVVKFLIKKKNALNDLRYGR TALIL 311		
QY	358	AASKNRVVCSSLLSYGADPTLLNCHNKSADIDAPT--QKLERLAYEPKGHSLQGAARE 415		
Db	312	AVCCGSASTVLLLEQINVDSSQDLSGQARETAVSHHHVICOQLSDYKEOMUKISSE 371		
QY	416	----ADVTRIKKHLSEMVNFKHPOTH---ETALHCAAA--PYPKRQIOICELLRLKGA 465		
Db	372	NSNPNVSRTR-----NKPRTHMVVEVDSMPAASSVKKPFGLRSKMGKWCRCPP 421		
QY	466	NINEKTEFTPLHVASEKAHNDVVEVYKHEAKVNALDNLGQTSLHRAAYCGHLOT CRL 525		
Db	422	CCRESGKS-----NVGTSGDHDD-----SAMKTLRSKMGR-----WCRHCFPC-- 459		
QY	526	LLSYGCDPNIIISLQGF TALQMGNENVOQLLQEGISLGNSEADROLLEAAKAGDVEYVKL 585		

RESULT 12


```
; Patent No. US20020077458A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, a
; FILE REFERENCE: PT018P1
; CURRENT APPLICATION NUMBER: US/09/835,788A
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/159,585
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/167,246
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-788A-12

Query Match          6.1%; Score 392; DB 10; Length 740;
Best Local Similarity 25.9%; Pred. No. 1.1e-19;
Matches 156; Conservative 71; Mismatches 207; Indels 168; Gaps 15;

QY 39 RALSASFGGLALLAGPG-LLRLALLAVAAARIMSGRCAGGGAAC----- 87
Db 249 QGLLVQEPDG-LMVATPAQTLTDTLDDLIAAVSTRVPTGSNSSQTECLTPESCQTT 306
QY 88 ASAAAEAVEPA-----ARELFEACRNGDVERVKRLVTPF-KVNSRSDTAGRK 132
Db 307 SNVASQSMPPVPSVDIDAHTESNHDLTALTACAGGHEELVSVLIARDAKIEHRDKG-- 364
QY 133 STPLHFAAGFRKDVVYLLQNGANVQARDGGL-IPHNACSFHGAHVNVNLLLRHGADP 191
Db 365 FTPLILAATAGHVGVVILLDKGDIQAQSEPTKDTPLSLACSGGROEVVDLLARANK 424
QY 192 NARDNNYTLPLHAAIKGKIDVICIVLQHGAEPTIRNTDGTALDLDAPSAAVLTGEYK 251
Db 425 EHRNVSQDYTPLSLAASGGYVNIKIILLNAGAE----- 456
QY 252 KDELLESARSGNEKMMALLTPLNVNCHASDGRKSTPLHLAGYNRVKIYVOLLLOHCAV 311
Db 457 -----INSRTGSKLGISPLMLAAMNGHVPYAVKLLDMDGSDI 492
QY 312 HAK-DKGLVPLHNACSYGHEVTELLVKKHGACVNAMDLMQFTPLHDEAAKSNRVEVCSLL 370
Db 493 NAQIETNRNTALTACFQGRAEVVSVLLDRKANVEHRAKTGLTPLMEASGGYAEVGRVL 552
QY 371 LSYGADPTLLNCHKNKAIDAPTPQLKERLAYEFKSHLSLQAAAREADVTRIKKHLSEMV 430
Db 553 LDKGAD-----VNAPPV----- 564
QY 431 NFKHPOTHETALHCAASPYPRKQICELLRLKGANINEKTKBELTPLHVASEKAHNDVV 490
Db 565 -----PSSROTALTIAADKGHYK---FCELLIHGAIIDVRKNGTPTLWLASNGGHFDVV 617
QY 491 EVVVKHEAKVNALDNIQTSLHRAAYCGHLQTCRLLS 528
Db 618 QLLVQAGADVDAADNRKKTPLMSAFRKGHVKVVOYLKVEVNFQPSDIECMRYIATITDKE 677
QY 529 -----YCCDPNIISLOGFTALMGNNVQOLLOE-GISLGNSEADQRLLEAAKAGDVETV 582
Db 678 LLKKHCOCVETIVKAKDQAAE-ANKNASILLKELDLEKRSRESRKQALAAKREKRKR 736
QY 583 KK 584
Db 737 KK 738

RESULT 15
US-09-947-199-2
```

```
; Sequence 2, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-947-199-2

Query Match          5.1%; Score 332.5; DB 10; Length 835;
Best Local Similarity 26.1%; Pred. No. 2.2e-15;
Matches 111; Conservative 60; Mismatches 124; Indels 131; Gaps 14;

QY 422 KKHLSEMVNFKHP-----QTHETALHCAASPYPRKQICELLRLKGANINEKTKBELTLP 477
Db 81 KSHIRTLMLKGLRPSRLTRNGFTALHLAV---YKDNAELITSLHSGADIQQVGYGGLTA 137
QY 478 LHVASEKAHNDVVVEVVKHEAKVNALDNIQTSLHRAAYCGHLQTCRLLSYGDNPILS 537
Db 138 LHATITAGHLEAADVLLQHGANNVQIDAVFTPLHIAAYYCHEQVTRLLKFGADVNV-- 195
QY 538 LOGFTALMGNNVQOLLOEGISLGNSEADQRLLEAAKAGDVETVKKLCTVQSVNCRDIE 597
Db 196 -----SGEVDG----- 201
QY 598 GRQSTPLHFAAGYNRVSVVEVYLQHG--ADVHAKDKGLVPLHNACSYGHEVAELLVKH 655
Db 202 -----RPLHLASAKGFNLIAKLLMEEGSKADVNQADNEDHVPFLHPCSRFGHHDIIVKYLQS 257
QY 656 GAVV--NVADLWKTPLHDEAAKGYEICKLLLO-HGADP-TTKNRDGT----- 701
Db 258 DLEVQPHVNVNIYGDTPHLACYNGKFEVAKETIIQISGTESLTKENIFSETAFHSACTYK 317
QY 702 PLDLVKDGDTDIODLLRGDAALLDAAKKGLARVKKLSPPDNVNCRTDQGRHS-TPLHLA 760
Db 318 SIDLVK-----FLLD-----QNVINIHQGRDGTGHLHA 347
QY 761 AGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAF 820
Db 348 CYHGHIRLVQFLDNGADMN-----LVACDPKSKSEKDEQTCMLM-----WAY 390
QY 821 TPLHEA 826
Db 391 EKGHDA 396

Search completed: February 12, 2003, 07:53:05
Job time : 31.4359 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 03:07:14 ; Search time 24.3761 Seconds
(without alignments)
4890.316 Million cell updates/sec

Title: US-09-843-159B-4
Perfect score: 6464
Sequence: 1 RCSARRGAAGGCGAQRGARY.....AYPEYLITYQIMRPEGMVDG 1240
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	895	13.8	1765	2 T42714	ankyrin 3, splice
2	895	13.8	1961	2 T42716	ankyrin 3, splice
3	892.5	13.8	4377	2 A55575	ankyrin 3, long sp
4	891	13.8	1943	2 T42713	ankyrin 3, splice
5	889	13.8	1940	2 T42715	ankyrin 3, splice
6	880.5	13.6	1856	2 B35049	ankyrin 1, erythro
7	880.5	13.6	1880	2 A35049	ankyrin 1, erythro
8	878.5	13.6	1881	1 SJHUK	ankyrin 1, erythro
9	877	13.6	3924	2 S37431	ankyrin 2, neurona
10	876	13.6	1848	2 S37771	ankyrin, erythrocy
11	874	13.5	1862	2 T49502	ankyrin - mouse
12	866.5	13.4	1549	2 T13940	ankyrin - fruit fl
13	858	13.3	2039	2 T15347	ankyrin-related un
14	657	10.2	791	2 T42691	hypothetical prote
15	563.5	8.7	934	1 H71374	probable ankyrin -
16	552	8.5	1411	2 S30355	alpha-latiroinsecto
17	507.5	7.9	1058	2 D82654	ankyrin-like prote
18	494.5	7.7	1062	2 T30255	inversin - mouse
19	489.5	7.6	1062	2 T14151	Inv protein - mous
20	485	7.5	1401	2 S11527	alpha-latirotoxin p
21	463	7.2	1188	2 T19352	hypothetical prote
22	452.5	7.0	426	2 AE2149	hypothetical prote
23	426	6.6	991	2 T25412	hypothetical prote
24	416.5	6.4	1184	2 T00253	gene Ankhzn protei
25	399	6.2	2584	2 T24158	hypothetical prote
26	399	6.2	2606	2 T24157	hypothetical prote
27	389.5	6.0	397	2 T46445	hypothetical prote
28	370	5.7	1031	2 T43458	hypothetical prote
29	350	5.4	557	2 T46507	hypothetical prote

30 348.5 5.4 606 2 AC2508 hypothetical prote
31 345 5.3 1423 1 T37275 death-associated p
32 343 5.3 1107 2 T21280 hypothetical prote
33 329.5 5.1 1435 2 T32930 hypothetical prote
34 321 5.0 1398 2 T21884 hypothetical prote
35 315.5 4.9 656 2 A34793 sex-determining pr
36 307 4.7 1004 2 A55142 myosin-light-chain
37 306 4.7 1016 2 T19006 ankyrin related pr
38 306 4.7 1322 2 A59288 myosin heavy chain
39 305 4.7 679 2 B45771 2-5A-dependent RNA
40 299.5 4.6 633 2 T27499 hypothetical prote
41 299 4.6 247 2 D84448 probable ankyrin l
42 297 4.6 1001 2 S30385 G9a protein - huma
43 297 4.6 1295 2 T21720 hypothetical prote
44 294.5 4.6 900 2 A42024 transcription fact
45 285.5 4.4 658 2 S68418 protein phosphatas

RESULT 1
T42714
ankyrin 3, splice form 2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42714
R:Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo
J. Cell Biol. 130, 313-330, 1995
A:Title:Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin ge
the repeat domain.
A:Reference number: 222237; MUID:95340633; PMID:7615634
A:Accession: T42714
A>Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: mRNA
A:Residues: 1-1765 <P>
A:Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605.1
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
A:Introns: 1587/1
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

Query Match 13.8%; Score 895; DB 2; Length 1765;
Best Local Similarity 30.3%; Pred. No. 1.6e-46;
Matches 294; Conservative 134; Mismatches 409; Indels 134; Gaps 25;

Qy 68 AVAAARIMSGRRRCAGCGGACASAAAAEVEPAARELFECACRNGDVERVKRLVTPKVNSRD 127
Db 10 AKPAHRKRGKGSNDANASYLRAARAGHLEKA----LDYIKNG-----VDVNICN 54
Qy 128 TAGRKSTPLHFAAGRGKDVVEYLLQNGANVQARDGGILPLHNACSFGEAEVWLLRH 187
Db 55 QNGLNA--LHLASKEGHVEVSELLQREANVDAATKKGNTALHLIASLAGAQAEVVKVLTN 112
Qy 188 GADPNARNWNWYTPHAAIKGIDVCIVLLOHGAEPITRNTDGTALDLA-----DPSA 242
Db 113 GANVNAQSQNGFTPLYMAAQENHLEVRFLDNGASQSLSLATEGFTPLAVALQQGHQDV 172
Qy 243 KAVLTGEYK---KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHAAAGYNRVK 299
Db 173 SLLLENDTKGVRLPALHIAARKDDTKAAALLQNDTNADVESKSGFTPLHIAAHYGNIN 232
Qy 300 IVQLLLQHGADVHAKDGLVPLHNACSYGHEVTELLVKHGACVAMDLWQFTPLHEAA 359
Db 233 VATLLNRAAAVDFTARNDDITPLHVASKRGANVMVKLLLDRCAGAKIDAKTRDGLTPLHCGA 292
Qy 360 SKNRVEVCSLLLSYGADFTLLNCHNKSALDAPTQPKERLAYEFKSHSLQAAAREADVT 419
Db 293 RSGHEQVVEMLLDRSA-PILSKTKN-----GLSPLHMAIQGD-- 328

ALIGNMENTS

Db 55 QNGLNA--LHLASKEGHEVWVSELOREANVDAATKGNLTALHIASLAGOAEVVKVLVTN 112
QY 188 GADPNARDNNWYTPLEHAAIKGIDVICVILQHGAEPTIRNTDGRALTDLA-----DPSA 242
Db 113 GANVNAOSONGFTPLYMAAQENHLEVVRLDNGASQSALATEDGFTPLAVALQOQHDOVV 172
QY 243 KAVLTGEYK---KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVK 299
Db 173 SLLENDTKGVRLPALHIAARKDDTKAAALLLONDNDVDSKSGFTPLHIAAHYGNIN 232
QY 300 IVOLLQHGADVHAKDGLVPLHNAACSYGHEVTELLVKGACVNMADLWQFTPLHEAA 359
Db 233 VATLLNRAAAVDFTARNITPLHVASKRGNANVKKLLDRGAKIDAKTRDGLTPLHCGA 292
QY 360 SKNRVECSLLSYGADPTLLNCHNSAIDLAPTPOLKERLAYEFKGHSLLOQAAREADVT 419
Db 293 RSHEQVEMVLENDRA--PILSKTKN-----GLSLPHMATOGD-- 328
QY 420 RIKKHLSLEMVNFH-----PQTHETALHCAASAPYKPKQICELLRLKGANINEKTK 472
Db 329 ----HLCVQGLLOHNVDPVDDVNDYLTALHVAACHGCHYK---VAKVLLDKKASPNKAL 381
QY 473 EFTLPLHVAASEKANDVVEVVVHKAQVNDLNGOTSLHRAAYCGHLOTCRLLSYGCD 532
Db 382 NGFTPLHIACKKRNIRVMELLKHGASIQAVTESGLTPIHVAAFMGHVNTVSQLMHHGAS 441
QY 533 PNIIISQGGTALOMG-----NENVOQLQEGISIGNSEADRG--LLEAAKAGDVETVKKL 585
Db 442 PNTNVRGETALHMAARSGOAEVVRVLDQGAQVEAKAKDDQTPHLISARLGKADIVQOL 501
QY 586 CTV-QSVNCRDIEGROSTPLHFAAGYNRVSVVEYLLQHGADVHAKDGLVPLHNAACSYG 644
Db 502 LOQASPNAAATTSG--YTPLHAAAREGHEDVAAFLLDHGASLSITTKKGTPLHVAAYK 559
QY 645 HYEVAELLVKGAVVNVADLWKTTPLEHAAAKGKYEICKLLLOHGADPTKKNRDNTPLD 704
Db 560 KLEVASLLQKASPDAAAGSGTPLHVAAHYDNQVALLLLDQGASPHAAAKNGYTPLH 619
QY 705 L-VKDGDTDIDOLL-----RGDAALDAAKGLCARVKKLSSPD-VNVCRTDQ 750
Db 620 IAAKNQMDIATSELLEGADANAVTQGTASVHLAQEGHVDVMVSLLSRANVNLSNKS 679
QY 751 GRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDAALLIKYNA 810
Db 680 G--LTPLHLAAQEDRVNAEVLVQGAHVDATQKMGYTPPLHVGCHYGNIKIVNFFLQHSA 737
QY 811 CVNATDKWATPLHAAQKORTOLCALLLAHGADPTLKNOEGTPLDLVS-----ADD 863
Db 738 KVNAKTNGYTAHQAAQOQGHTHIINVLLQNNASPNELTVNGNTALAIARRLGYISVVD 797
QY 864 VSALLTAAMPSPALPSYK---POVLN-----GVRSPGAT-----AD 897
Db 798 LKVVTEIINTVITTEKHKNVPEVTHNEVLDMSDDEVRKASAPKLSGGEYSIDGEGED 857
QY 898 ALSS-----GP-----SSPSSLSAASSLDNLGSEFSELSVVVS-----SSOTEGA 937
Db 858 ATGTDTKYLGPDQLKELGDDSLPAEGYVGFSLGARSASLRSFSSDRSYTLNRSYARDS 917
QY 938 SSLEKKEVPGVDSITQFVNIGLEHLMDFEREQITLDVL-----VENGHKELKEIGI 991
Db 918 MMEIFELVPSKEGHLT--FTREFSDSLRH--YSWAADTLDNVNLVSPVHSGF--LVSFV 973
QY 992 NAYG-----HRHKLI-----KGVRLISQOQGLNPILTINTSGSGTILIDL 1033
Db 974 DARGSGMRGRHGMRIIPRCKTAPTITCRVLRKHKLANPPVWEGEGSLASRLVEMG 1033
QY 1034 PDKQFOS--VEEMQSTVREHROGGHAGGTIFNRYNLIKIOKVNKKLWE--RYTHRRKEV 1090
Db 1034 PAGAQLGLPVIVETP-----HFGSMRGKRELIIVLRSNGETWKEHQFDSKNED 1083
QY 1091 SE-----ENHNANERMLPHGSPFVNAIHKGFDERHAYIGGMFAGAGIYFAENS---SKS 1142

Db 1084 AELLNGMDELDSPLEEL---GTXKRICRIITKDFPO-----YFAVVSRIKQES 1127
QY 1143 NOYVYGIGGG 1152
Db 1128 NQ--IGPEGG 1135
RESULT 5
T42715
ankyrin 3, splice form 3 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42715
R:Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo
J. Cell Biol. 130, 313-330, 1995
A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A:Reference number: 222237; MUID:95340633; PMID:7615634
A:Accession: T42715
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1940 <PEP>
A:Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDN:AA01604.1
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
A:Introns: 834/1
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

Query Match 13.8%; Score 889; DB 2; Length 1940;
Best Local Similarity 27.9%; Pred. No. 4.4e-46;
Matches 337; Conservative 168; Mismatches 496; Indels 206; Gaps 38;

QY 68 AVAAARIMSRRCAGGACACASAAAEVPAARELFPACRGDVERKRLVTEKVNRSR 127
Db 10 AKPAHKKRKGKSDANASYLRAARAGHLEKA---LDYIKNG-----VDVNICN 54
QY 128 TAGRKSTPLHFAAGFGKDVVEYLLQNGANVQARDGGLIPLHNACSFHGAHVNNLLRH 187
Db 55 QNGLNA--LHLASKEGHEVWVSELOREANVDAATKGNLTALHIASLAGOAEVVKVLVTN 112
QY 188 GADPNARDNNWYTPLEHAAIKGIDVICVILQHGAEPTIRNTDGRALTDLA-----DPSA 242
Db 113 GANVNAOSONGFTPLYMAAQENHLEVVRLDNGASQSALATEDGFTPLAVALQOQHDOVV 172
QY 243 KAVLTGEYK---KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVK 299
Db 173 SLLENDTKGVRLPALHIAARKDDTKAAALLLONDNDVDSKSGFTPLHIAAHYGNIN 232
QY 300 IVOLLQHGADVHAKDGLVPLHNAACSYGHEVTELLVKGACVNMADLWQFTPLHEAA 359
Db 233 VATLLNRAAAVDFTARNITPLHVASKRGNANVKKLLDRGAKIDAKTRDGLTPLHCGA 292
QY 360 SKNRVECSLLSYGADPTLLNCHNSAIDLAPTPOLKERLAYEFKGHSLLOQAAREADVT 419
Db 293 RSHEQVEMVLENDRA--PILSKTKN-----GLSLPHMATOGD-- 328
QY 420 RIKKHLSLEMVNFH-----PQTHETALHCAASAPYKPKQICELLRLKGANINEKTK 472
Db 329 ----HLCVQGLLOHNVDPVDDVNDYLTALHVAACHGCHYK---VAKVLLDKKASPNKAL 381
QY 473 EFTLPLHVAASEKANDVVEVVVHKAQVNDLNGOTSLHRAAYCGHLOTCRLLSYGCD 532
Db 382 NGFTPLHIACKKRNIRVMELLKHGASIQAVTESGLTPIHVAAFMGHVNTVSQLMHHGAS 441
QY 533 PNIIISQGGTALOMG-----NENVOQLQEGISIGNSEADRG--LLEAAKAGDVETVKKL 585
Db 442 PNTNVRGETALHMAARSGOAEVVRVLDQGAQVEAKAKDDQTPHLISARLGKADIVQOL 501
QY 586 CTV-QSVNCRDIEGROSTPLHFAAGYNRVSVVEYLLQHGADVHAKDGLVPLHNAACSYG 644

```

Db 502 LOOGASPNATTS--YTPHLHAAREGHEDVAFLLDHGCASISITTKGFTPLHVAAYG 559
Qy 645 HYEVALLKHGAVNVADLWFTPLHEAAAGKYEICKLLLOHGADPTKKRDCNTPLD 704
Db 560 KLEVASLLLOKSPDAACKSGLTPLHVAHYDNOKVALLLDQAGSPHAAKNGYTPPLH 619
Qy 705 L-VKGDGTDIDQLL-----RGDAALLDAKAGKGLARVKKLSPD-NVNCRTDQ 750
Db 620 IAAKKNMIDATSLLEYGADANAVTQGTASVHLAAQEGHDMVSLLSRNVNLSNKS 679
Qy 751 GRISTPLHAAAGYNNLEVEAYLLOHGADVNAQDKGLPLHNAASVGHVDAALLIKYA 810
Db 680 G--LTPHLHAAQDRVNAEVLVNOQAHVDAQTKGYTPLHVGCHYKIKVNFLLQHS 737
Qy 811 CVNATDKWFTPLHEAAQGRQTQCALLAHGAADPTLKKQEGQTPLDLVS-----ADD 863
Db 738 KVNAKTKNGYALHQAQOQHIIINVLONNAPNELTVNGTALAIAARRLGYISVVD 797
Qy 864 VSALLTAAPPSALPSCYK---PQVLNGVRSPG-----ATADALS 900
Db 798 LKVVTEIIMTTTITEKHKNMNPETMNEVLDMSDDEGDKCTWFKIPKQGEVLVKSFDAIT 857
Qy 901 S-----GP-----SPSSLSAASLDNLGSELSVVS-----SSGTECASSL 940
Db 858 GDTXYGLPODLKELGDDSLPAGYVGFSLGARSASLSRFSFSDRSYTLNRSYARDMMI 917
Qy 941 EKKEVPGVDFISITQFVRLNGLEHLMDFEREQITLDVL-----VEMGHKELKEIGINAY 994
Db 918 EELLVPSKEQHLT-FTRFEDSDSLR-YSWAADTLDNVNLVSSPVHSGF--LVSEMVDA 973
Qy 995 G-----HRHKL-----KGVRLISGGQGLNPYTLTNVSGSGTILIDLSPDD 1036
Db 974 GSGMRGSRHGMRIIPPRKCTAPTITCRVLKRHLANPPMVEGEGLASRLVEMGPAG 1033
Qy 1037 KEFQS-VEEEMQSTVREHSDGAGGIFNRYNLTKQVCNKKLWE--RYTHRRKEVSP- 1092
Db 1034 AQPLGVVIEP-----HFGSMRGERELIVLRSENGETWKEHGFDSKNEIDLAE 1083
Qy 1093 -----ENHNHANRMLFHGSPFVNAIHKGFDERHAYIGMFGAGIYFAENS---SKSNQY 1145
Db 1084 LNMDELDSPEL---GTRICRIITKDFPQ-----YFAVVSRIKQESNQ- 1126
Qy 1146 VYIGGG 1152
Db 1127 -IGPEGG 1132

RESULT 6
B35049
N: ankyrin 1, erythrocyte splice form 3 - human
N: Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N: Contains: ankyrin 2.2, erythrocyte
C: Species: Homo sapiens (man)
C: Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
C: Accession: B35049
R: Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, M.C.; K
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A: Title: cDNA sequence for human erythrocyte ankyrin.
A: Reference number: A35049; MUID: 90175370; PMID: 1689849
A: Accession: B35049
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-1856 <LAM>
C: Genetics:
A: Gene: GDB:ANK1; ANK
A: Cross-references: GDB:118737; OMIM:182900
A: Map position: 8p11.2-8p11.2
C: Superfamily: ankyrin; ankyrin repeat homology
C: Keywords: alternative splicing
F: 2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MAT>
F: 2-1513,1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F: 44-76/Domain: ankyrin repeat homology <AN01>
F: 77-109/Domain: ankyrin repeat homology <AN02>

```

```

F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>
F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match 13.68; Score 880.5; DB 2; Length 1856;
Best Local Similarity 26.08; Pred. No. 1.4e-45;
Matches 294; Conservative 156; Mismatches 374; Indels 307; Gaps 29;

Qy 105 ACNRGSDVERKRLVTPKVNRSRDYAGRKSTPLHFAAGFGRKDVVEYLQNGANGYQAARDG 164
Db 52 ASKEGHVKVWVVELLHKE-IILETTTKGNTALHIAALAGQDEVVRELVNVCANVAQSOK 110
Qy 165 GLIPLHNACSPGHARVWMLLRHGHADPNARDNNWYTPHLEAAIKGKIDVCTVLQHGAE 224
Db 111 GFTPLMAAENHLEVVVKFLLENGANQVATEDGFTPLAVALQOQHENVVAHLINCYTKG 170
Qy 225 TI-----RNTDGRALDLADSAKAVLTGEYKDELLESARSGNEERKMALLPLN 275
Db 171 KVRPLALHIAARNDTRTA-----AVLLNDPNPDL--SKTG-----FTPLH 211
Qy 276 VNCHASD-----GRKS-----TPHLAAGYNRYKIVOLLQHGADVHAKDKG 317
Db 212 IAAHYENLNVAGLLNRGSSVNFTPONGITPLHIASRRGNVIMVRLLDLDRGAQIETKTD 271
Qy 318 DLVPLHNACSVGHYEVTTELLVKGACVNAWDLWQFTPLHEAASKNRVEVCILLSYGD 377
Db 272 ELTPLHCAARNGHVRISBELLDHGAPIOAKTKNGLUSPTHMAAQGDHDCVRLLLQYDAE- 330
Qy 378 TLLNCHNKSADLAPTQPKERLAYEFGHSLQAAREADVTRIKKHLSELVNFKHPQT 437
Db 331 -----ID-----DIT-----LD 337
Qy 438 HETALHCAAAPYKPKKQICELLRRKGANINEKTFEFTPLHVAASEKAHNDVVEVVKHE 497
Db 338 HLTPLHVA--HCGHHRVAKVLLDKGAKPNSRALNGFTPLHIAACKNHRVMEILLKTG 394
Qy 498 AKVNALDNLGQTSILHRAAYCGHLQTCRLLLSYGCDPNIISLOGFTALONG-----NENVQ 552
Db 395 ASIDAVTESGLTPLHVASFMGHLPVTKNLLQRGASPNVSKVETPLHMAARAGHTEYAK 454
Qy 553 QLLQE-----GISLGNSEADROLLE-----AAKAG 577
Db 455 YLLQNKAKVNAKAKDDQTPHCAARIGHNTNMVKKLLENNANPNLATTAGTPTLHIAAREG 514
Qy 578 DVETVKLCTVQ-SYNCRDIEGRQSTPLHFAAGYNRVSVVEYLQHGADVHAKDKGLVP 636
Db 515 HVETVLALLEKEASQACMTKKG--FTPLHVAAYKVKRVAELLERDAHPNAAAGNGLTP 572
Qy 637 LHNACSVGHYEVAELLVKGAVNVADLWK-FTPLHEAASKYIEICKLLQHGADPTKK 695
Db 573 LHVAVHHNNLDIVKLLPLRGSGSPH-SPAWNGYTPHIAAKQOVAVRSLLQYGSANAE 631
Qy 696 NRDGNTPLDL-VKGDGTDIDQLRGDAALLDAAKGLARVKKLSPPDNVNCRDTPGRHS 754
Db 632 SVOGVTPLHAAQEGHAEWVALL-----LSKQANGNLGNKSG--L 669

```


Db 603 AAHYDNOKVALLLEKASPHATAKNGYTPPLHIAAKKNQMOIASTLLNNGYGAETNIVTKQG 662
Qy 700 NTPLDLV-KDGDITDQDLRLGDAALLDAAKKGGCLARVKKLSSPDNVNCRDQGRHSTPLH 758
Db 663 VTPLHLASOECHTDMVTLKDKGNIHMTKSL-----TSLH 700
Qy 759 LAAGYNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNATDKW 818
Db 701 LAAGEDKYNVADIITKHGADQDAHTKLGYPPLIVACHYGNVMVNFLLKOGANNAKTGN 760
Qy 819 AFTPLHEAAOKGRTQLCALLAHGADPTLKNQEGOTPLDL-----VSADDVSALLTAAM 872
Db 761 GYTPLHQAQOQHHTIINVLQHGAKGNATTANGNTALAIKRLGYISVDPLKVVTEEV 820
Qy 873 PPSALPCYKQVLNGVRSPGATADALSSGSPSSLSAASLDNLGSGFSEL----- 925
Db 821 TTTTITTEK-----HKLNVPTMETEVL-----DYSDEEGDDTMTCDGGEYLRPEDLK 868
Qy 926 ----SSVVSSGTEGASLEKKEVPGVDFSTQFV--RNGLGLEHLMDFEREQITLVLV 979
Db 869 ELGDDSLPSSQFLDGMNLYRYSLEGGRSDSLSRFSFSDRSHTLSHAS--YLRDSAMVMDSV 926
Qy 980 EMGHKELKEIGINAYGHRHKLKIGVERL 1007
Db 927 VIPSHQVSTLAKEARNYRLSWGTENL 954
RESULT 10
S3771
ankyrin, erythrocyte - mouse
C:Species: Mus musculus (house mouse)
C>Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C:Accession: S3771
R: Birkenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.
J. Biol. Chem. 268, 9533-9540, 1993
A: Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found an
A: Reference number: S3771; MUID: 93252825; PMID: 8486643
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-1848 <BIR>
A: Cross-references: EMBL:X59063; NID:g311816; PIDN:CA448801.1; PID:g311817
C: Superfamily: ankyrin; ankyrin repeat homology
C: Keywords: alternative splicing
F: 48-80/Domain: ankyrin repeat homology <AN01>
F: 81-113/Domain: ankyrin repeat homology <AN02>
F: 114-146/Domain: ankyrin repeat homology <AN03>
F: 147-175/Domain: ankyrin repeat homology <AN04>
F: 176-208/Domain: ankyrin repeat homology <AN05>
F: 209-241/Domain: ankyrin repeat homology <AN06>
F: 242-274/Domain: ankyrin repeat homology <AN07>
F: 275-307/Domain: ankyrin repeat homology <AN08>
F: 308-340/Domain: ankyrin repeat homology <AN09>
F: 341-373/Domain: ankyrin repeat homology <AN10>
F: 374-406/Domain: ankyrin repeat homology <AN11>
F: 407-439/Domain: ankyrin repeat homology <AN12>
F: 440-472/Domain: ankyrin repeat homology <AN13>
F: 473-505/Domain: ankyrin repeat homology <AN14>
F: 506-538/Domain: ankyrin repeat homology <AN15>
F: 539-571/Domain: ankyrin repeat homology <AN16>
F: 572-604/Domain: ankyrin repeat homology <AN17>
F: 605-637/Domain: ankyrin repeat homology <AN18>
F: 638-670/Domain: ankyrin repeat homology <AN19>
F: 671-703/Domain: ankyrin repeat homology <AN20>
F: 704-736/Domain: ankyrin repeat homology <AN21>
F: 737-769/Domain: ankyrin repeat homology <AN22>
F: 770-802/Domain: ankyrin repeat homology <AN23>
Query Match 13.6%; Score 876; Db 2; Length 1848;
Best Local Similarity 29.7%; Pred. No. 2.6e-45;
Matches 250; Conservative 132; Mismatches 281; Indels 178; Gaps 21;
Qy 105 ACRNGDVERKRLVTPKVNRSRTAGRKSTPLHFAAGGRKDVVEYLLQNGANVQARDG 164

Db 56 ASKEGHVKNVVELLHKE-IILETTTKGNTALHIAALAQDEWRELNVNGANVAQSOK 114
Qy 165 GLIPLHNACSGHAEEVWNLLRHGADPNARDWNWYTPPLHEAIAKIDVICIVLLOHGAEP 224
Db 115 GFTPLYMAAQENHLEEVVKFLLENGANQVATEDGFTPLAVALQOQHENVVAHLINVTGK 174
Qy 225 TI-----RNTDGRATDLADPSAKAVLTGEYKKBELLESARSNEKEMMALITPLN 275
Db 175 KVRPLALHIAARNDDTRTA-----AVLLQNDPNPDVL--SKTG-----FTPLH 215
Qy 276 VNCHASDGRKS-----TPHLAAGYNRVKIVOLLQLOHGADVHAKDKG 317
Db 216 TAAHYENLVNAQALLNRGASVNFTPQNGITPLHIAASRGVIMVRLLLDRGAQIETRTKD 275
Qy 318 DLVPLHNACSYGHVEVTELLVKHGACVAMDLWQFTPLHEAAKNRVVCSSLLSYGADP 377
Db 276 ELTPLHCAARNGHVRISEILDDHGAPIQAKYKNGUSPIHMAAQGDHLLDCLVLLQTNAB- 334
Qy 378 TLLNCHNKSAIDLAPTQPKERLAYEFKGSLLQAAAREADVTRIKKHLSEMFNEKHPOT 437
Db 335 -----ID-----DIT-----LD 341
Qy 438 HETALHCAAASPYPKRKOICELELLRKGANINEKTEFTPLTHVASEKAHNDVVEVVKHE 497
Db 342 HLTPPLHVAA--HCGHHRVAKVLLDKGAPNSRALNGETPLHIAACKNHIRVMELLKTKG 398
Qy 498 AKVNALDNLGOTSLHRAAYCGHLOTCRLLLSYGCDPNIISLOGFTALQM----GNENVQO 553
Db 399 ASIDAVTESGLTPLHVASFMGHLPITVKNLLQRGASPNVSNVKVETPLHMAARAGHTEVA 458
Qy 554 LLQEGISLGNSEAD-----ROLLE-----AAKAG 577
Db 459 YLLONKAKANAKAKDDQTPHLCAARIGHTGVKVLLENGASPNLATATTAGTPLHTAAREG 518
Qy 578 DVEYVKILCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVP 636
Db 519 HVDYALALLEKEASQACMTKG--FTPLHVAAYKGVRLAEELLEDAHPNAAGNGLTP 576
Qy 637 LHNACSYGHEVAELLVKHGAVVNVADLWK-FTPLHEAAAKGYEICKLLLOHGADPYK 695
Db 577 LHVAVHHNLLDIVKLLLPRLGGSPH-SPAWNGYTPPLHIAAKQOIIEVARSLLOQYGSANAE 635
Qy 696 NRDGNTPLDL-VKGDGDTDQDLRGDAALLDAAKKGLARVKKLSSPDNVNCRDQGRHS 754
Db 636 SVQGYTPPLHAAQEGHTEWALL-----LSQANGNLGNKSG--L 673
Qy 755 TPLHLAAGYNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNA 814
Db 674 TPLHLVSOEGHVPVADVLKIKGVTVDAITRMGYTPLHVASHYGNIKLVKFLQHQADVNA 733
Qy 815 TDKWAFTPLHEAAOKGRTQLCALLAHGADPTLKNQEGOTPLDL-----VSADDVSALL 868
Db 734 KTKLGSPLHQAQOQHHTDITVLLKNGASPNVSSNGTTPPLAIAKRLGYISVTVLKV 793
Qy 869 T 869
Db 794 T 794
RESULT 11
I49502
ankyrin - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C:Accession: I49502
R: White, R.A.; Birkenmeier, C.S.; Peters, L.L.; Barker, J.E.; Lux, S.E.
Mamm. Genome 3, 281-285, 1992
A: Title: Murine erythrocyte ankyrin cDNA: Highly conserved regions of the regulatory
A: Reference number: I49502; MUID: 92345717; PMID: 1386265
A: Accession: I49502
A: Status: preliminary
A: Molecule type: mRNA

A:Residues: 1-1862 <RES>
A:Cross-References: GB:M84756; NID:g191939; PIDN:AAA37236.1; PID:g191940
C:Genetics:
A:Gene: Ank-1
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:40-72/Domain: ankyrin repeat homology <AN01>
F:73-105/Domain: ankyrin repeat homology <AN02>
F:106-138/Domain: ankyrin repeat homology <AN03>
F:139-167/Domain: ankyrin repeat homology <AN04>
F:168-200/Domain: ankyrin repeat homology <AN05>
F:201-233/Domain: ankyrin repeat homology <AN06>
F:234-266/Domain: ankyrin repeat homology <AN07>
F:267-299/Domain: ankyrin repeat homology <AN08>
F:300-332/Domain: ankyrin repeat homology <AN09>
F:333-365/Domain: ankyrin repeat homology <AN10>
F:366-398/Domain: ankyrin repeat homology <AN11>
F:399-431/Domain: ankyrin repeat homology <AN12>
F:432-464/Domain: ankyrin repeat homology <AN13>
F:465-497/Domain: ankyrin repeat homology <AN14>
F:498-530/Domain: ankyrin repeat homology <AN15>
F:531-563/Domain: ankyrin repeat homology <AN16>
F:564-596/Domain: ankyrin repeat homology <AN17>
F:597-629/Domain: ankyrin repeat homology <AN18>
F:630-662/Domain: ankyrin repeat homology <AN19>
F:663-695/Domain: ankyrin repeat homology <AN20>
F:696-728/Domain: ankyrin repeat homology <AN21>
F:729-761/Domain: ankyrin repeat homology <AN22>
F:762-794/Domain: ankyrin repeat homology <AN23>

Query Match 13.5%; Score 874; DB 2; Length 1862;
Best Local Similarity 29.7%; Pred. No. 3.5e-45;
Matches 250; Conservative 132; Mismatches 281; Indels 178; Gaps 21;

QY 105 ACRCNGDVERVKRLVTEPKVNSRDTAGKSTPLHFAAGFGKDWVEYLLONGANVQARDG 164
Db ASKEGHVKMVELLHKE--IILETTTKKGTALHIAALAGODEYVELVNGYANVNAQSOK 106

QY 165 GLIPLNACSFHAEVYVLLRHGADPNARDNNWYTPLEHAALKGKIDVICVILQHGAPP 224
Db GTPLYMAAQENHLEVVKVFLLENGANQNVATEDGFTPLAVALQOQHENVVAHLINVTGK 166

QY 225 TI-----RNTDORTALDADPSAKAVITGEYKKDELLESARGSENEKMMALLTPN 275
Db KVRPLPALHTAARNDDRTA-----AVLLQNDPNPDVL--SKTG-----FTPLH 207

QY 276 VNCHASDGRKS-----TPLHLAGYNRVKIVQLLLQHGADVHAKDG 317
Db IAHYENLVNQALLNRGASVNFPTONGITPLHIAARRGNVIMVRLLLDRGAQIETRKD 267

QY 318 DLVPLNACSYGHEVTELLVKHGACVNAIDLWQFTPLHEAASKNRVEVCSLLLSYGADP 377
Db EUTPLHCAARNHVRISSEILLDHGAPLOAKTNGLSPIHMAAQGDHDCVRLLLQYNAE- 326

QY 378 TLLNCHNKSAIDLAPTPOLKERLAYEFKGHSLLOQAAREADVTRIKHLSLEVMVNFKHPT 437
Db ID-----ID-----DIT-----LD 333

QY 438 HETALHCAASYPKPKQICEILLRLKGANINEKTEFLPLHVASEKAHNDVVEVVVKE 497
Db HLTPLHVA--HCGHRRVAKVLLDKGAPNSRALNGFTPLHIACKKNHIRMVLELLKRG 390

QY 498 AKVNALNDLGOTSLHRAAYCGHLOTCLRLLSYCGCDPNILSLOGTALOM---GNENVOQ 553
Db ASIDAVTESGLTPLHVASFPGHLPVKNLLORGASPNVSNVETPLHMAARAGHTEVAK 450

QY 554 LLOEGTSLNSAD-----ROLLE-----AAKAG 577
Db YLLONKAKANAKAKDDOTPLHCAARTGHTGMVKLLLENGASPNLATTAGHTPLHTAAREG 510

QY 578 DVETVKKLTQV--SYNCRDIEGRQSTPLHFAAGYNEVSVVYLLQHGADVHAKDGGLVP 636
Db HVDTALLLEKASQACMTKKG--FTPLHVAAYKGVKRLAELELLEHDAHPNAAKGKGLTP 568

QY 637 LHNACSYGHEVYAEELLVKHGAVVNVADLMK-FTPLHEAAAKGYEICKLLQLQHGADPTKK 695
Db LHVAVHHNLLDIVKLLPRGSGPH-SPANNGYTPPLHIAAKQKQIEVARSLQYGGSANAE 627

QY 696 NRDGNTPDL-VKGDGTDIDQLLRGDAALLDAAKCKLARVKKLSPPDNVNCRDTOGRHS 754
Db SVQGVTPPLHAAQEGHTEMVALL-----LSQANGNLGNKSG--L 665

QY 755 TPLHLAGYNNLEVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNA 814
Db TPLHLVSEQHVLVADVLIKHGVTVDATTRMGYTPPLHVASHYGNIKLVKFLQHQADVNA 725

QY 815 TDKWAFTPLHEAAQKQRTQCALLLAHGADPTLKNOEGOTPLDL-----VSADYSALL 868
Db KTKLGYSPHLHQAQOQHTDIVTLLKNGASPNESVSSNGTTPPLAIAKRLGLVISVTDVLKV 785

QY 869 T 869
Db 786 T 786

RESULT 12
TI3940
ankyrin - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C:Accession: TI3940
R:Dubreuil, R.R.; Yu, J.
Proc. Natl. Acad. Sci. U.S.A. 91, 10285-10289, 1994
A:Title: Ankyrin and beta-spectrin accumulate independently of alpha-spectrin in Dros
A:Reference number: Z17820; MUID:95024098; PMID:7937942
A:Accession: TI3940
A>Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1549 <DUB>
A:Cross-References: EMBL:L35601; NID:g557083; PID:g557084; PIDN:AAAC37208.1
C:Genetics:
A:Cross-References: FlyBase:FBgn0011747

Query Match 13.4%; Score 866.5; DB 2; Length 1549;
Best Local Similarity 27.4%; Pred. No. 7.5e-45;
Matches 311; Conservative 173; Mismatches 492; Indels 159; Gaps 34;

QY 98 AARELFEACRNGDVERVKRLV---TPEKVNSRDTAGKSTPLHFAAGFGKDWVEYLLQN 154
Db ATISFLRAARSGDIKKVMDFLDCGEISDINSCNANGLNA--LHLAKDGVVDICCELLRR 95

QY 155 GANVQARDGGIPLPLNACSFHAEVYVLLRHGADPNARDNNWYTPLEHAATKGDV 214
Db GIKIDNATKKGNTALHTASLAGQHDVINQLLYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLQHGAEPTTNTDGRALDLA-----DPSAKAVITGEYK---KDELLESARSNEEK 266
Db RTLLANGANPSLSTEDGFTPLAVAMQOQGHDKIVAVLELNDVVRGKVRPLHIAAKNDVN 215

QY 267 MMALTPLVNCHASDGRKSTPLHLAGYNRVKIVQLLLQHGADVHAKDGDLVPLHNA 326
Db AKLLQLHDPNADIVSKSGFTPLHIAAHYGNVDIATLLNKNKADVNVVAKHNTPPLHVAC 275

QY 327 SYGHEVTELLVKHGACVNAIDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN-K 385
Db KWGLSLCTLLLCRGAKIDAATRDGLTPLHCASRSHVEVIKHLQONA-PILTCKNGL 334

QY 386 SAIDLAPTOLKERLAYEFKGHSLLOQAAREADVTRIKHLSLE-----MVNF 432
Db SALHMAAQE-----HDEAAHILLDNKAPDEVTVDTLHVAACHGKVKVAKLLDY 388

QY 433 K-HPOTHE---TALHCAASYPKPKQICEILLRLKGANINEKTEFLPLHVASEKAH 487
Db KANPNARALNGFTPLHIACKK---NRKVVELLIKGANIGATTESGLTPLHVASFPGCI 445

QY 488 DVVEVVVHKEAKVNALDNLGOTSLHRAAYCGHLOTCLRLLSYCGCDPNILSLOGFTAL--- 544

Db 446 NIVYLLQHEASADLPTIRGETPLHLAARANQADIIRILLR-SAKVDAIVREGQTPHLHVA 504
QY 545 -QMGEN-VQOLLEGLISGNSEADR--QLLEAAKAGDVETVKKLCIVQSVNCRDIERGQ 600
Db 505 SRLGNIIMILLQHGAEINNAQSNKYSAHIAAKEGOENIVOVLLANGAEN-NAVTKKG 563
QY 601 STPLHFAAGNRVSVYVLLQHGADVHAKDKGGLVPLHNAACSYGHEVAEALLVKHGAVERN 660
Db 564 FTPLHLACKYGVQNVQVILLONGASIDFQGNVDVPLHVAHYNNPSIVELLELANGSSPN 623
QY 661 VADLWKFTPLHEAAAKGYEICKLLQHGADPTKKNRDGNTPDL-VKDGDTDIDQDLRG 719
Db 624 LCARNQOCATHIACKKNVLEIAMQLLQHGADVNIISKSGFPLHAAQGGNVDMVQLLLE 683
QY 720 DAALLDRAKKCLARVKLLSPDNVNCRTDGTGRHSTPLHLAAGYNNLEVAEYLLQHGADV 779
Db 684 YGVLSAAAKNGL-----TFLHVAQAGHVLVSGIILEHGANI 720
QY 780 NAQDKGGLIPLHNAASYGHVDAALLIKYNACVNATDKWFTPLHEAAQKRTOLCALLL 839
Db 721 SERTNGYTPLHMAAHYGHLDLVKFFIENDADIEMSSNIGYTPLHQAQOQHIMIINLLL 780
QY 840 AHGADPTLKNQDGTPLDVSADVSALLTAAMPSPALPSCYKPOVLNG-----V 889
Db 781 RHKANPNALTKDGN TALHASN-----LGYVTVMESLKIVTSTSVINSNIGATEEKLK 834
QY 890 RSPCATDALSSGSPSSLSAAS-----SLDNLSGSFSELSVSSVSSGTE-----CAS 938
Db 835 MTPELMOTELLSDDSCDLDLHNYKYMATDLDLANYQDQKNFTTNTDHLTDVS 894
QY 939 SLEKKEVPGVDFSTQFVRNLGLEHMD--TFEREQITLDLVEMGHKELKEIGINAYGH 996
Db 895 VLNKEILPNEMSCIELTE---IGHKPDNVVIARSQVHLGFLVSF---LVDARGSMRGY 948
QY 997 RHKLKIGVE-----RLISGOGLNPLYLTNTSGSGLTILDLSPDDKEFQS-- 1041
Db 949 RHNGVRIITVPPKACAEPTITRCYVKVQPVVPPPLMEGEALVSKILEMSPVDMGFLSPI 1008
QY 1042 -VESEMOSTVREHRDGGHAGIFRNYNLLKQVCNKKLWERYTHRRKEVSEENHNHANE 1100
Db 1009 TLEVPHYGTLRK-----NEREII-TLRSNDGESWREH-NLYKDDITGEDINQFEE 1055
QY 1101 RMLFHGSPFVNAIHKGPDERHAYIGGMFGAGIVFAENSSKSNQVYVIG--GGT 1153
Db 1056 ---FHSDDRIV-IVTQNPVPHFAVV-----SVRQEVHVGPDGGT 1092
RESULT 13
T15347
ankyrin-related unc-44 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Aug-2002
C:Accession: T15347; T15346; T15344; T15345; A57282; B57282; C57282
R:Gattung, S.
submitted to the EMBL Data Library, February 1996
A:Description: The sequence of C. elegans cosmid B0350.
A:Reference number: Z18332
A:Accession: T15347
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-2039 <GAT>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208873; PIDN:AAA93443.1
A:Accession: T15346
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-1000, 'SKLOHRT',1002-1718, 'KWEELNRL',1727,'D','ES',1942,'PSPAQRS',1950,'IVA
32','S',2034-2035,'GSPTRSRVEPEEHRHSHOHEGST' <GA2>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208875; PIDN:AAA93445.1
A:Accession: T15344
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-1718, 'KW',1903-1905,'NRLADESSPS',1916-1917,'QRSTIVAESTSEQVPE',1934-1935,'B

<GA3>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208874; PIDN:AAA93444.1
A:Accession: T15345
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-1718,'KWEELNRL',1727,'D','ES',1942,'PSPAQRS',1950,'IVAES',1956-1957,'EQ
PTRSRVEPEEHRHSHOHEGST' <GA4>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208876; PIDN:AAA93446.1
R:Otsuka, A.J.; Franco, R.; Yang, B.; Shim, K.H.; Tang, L.Z.; Zhang, Y.Y.; Boontrakul
J. Cell Biol. 129, 1081-1092, 1995
A:Title: An ankyrin-related gene (unc-44) is necessary for proper axonal guidance in
A:Reference number: A57282; MUID:95263663; PMID:7744957
A:Accession: A57282
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-852,'GGG',856-1000,'SKLOHRT',1002-1319,'IG',1322-1595,'DA',1598-1718,'K
'SHRED',2007-2008,'TI',2011,2017,'TT',2020-2022,'SHIS' <OTS>
A:Cross-references: GB:U21734; NID:g790607; PIDN:AAA85854.1; PID:g790608
A:Accession: B57282
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 831-852,'GGG',856-1319,'IG',1322-1595,'DA',1598-1718,'KWEELNRL',1727,1799
V',1945-1947,'VT',1950,'SH',1975,'SESP',1980-1981,'SPTRSRVEPEEHRHS',1984-1985,'EDHEGS
A:Cross-references: GB:U21731
A:Accession: C57282
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 194,'E',196,'I',198,'DC',201,'G',409,'AV',412,'Q',414,824,'S',826,'ER',82
4,'TIV',1828,'ESTS',836,'QV',1834-1935,'EOS',1939,'ESES',1944,'REDDGFTIVTT',
A:Cross-references: GB:U21732; NID:g790603; PIDN:AAA85853.1; PID:g790604
C:Genetics:
A:Gene: CESP:unc-44
A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1045/3; 1660/1; 1719/2; 1
C:Superfamily: ankyrin; ankyrin repeat homology
F:164-192/Domain: ankyrin repeat homology <ANQ4>
F:358-390/Domain: ankyrin repeat homology <ANI>
F:391-423/Domain: ankyrin repeat homology <ANL>
Query Match 13.3%; Score 858; DB 2; Length 2039;
Best Local Similarity 26.3%; Pred. No. 3.9e-44;
Matches 319; Conservative 163; Mismatches 443; Indels 286; Gaps 39;
QY 91 AAEEVPEAARE---LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRK 145
Db 22 APAAPEPEAGESASFLRAARAGDLEKLELRAGCTDINTSNANGLS--LHLASKEGHS 79
QY 146 DVVEYLLONGANVQARDGGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNWYTPLHEA 205
Db 80 EYVRELIRQAQVDAATRKNQALTALHIASLAGOSLIVTILVENGANVNVQSVNGFTPLYMA 139
QY 206 ATKGKIDVICVILQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKDELLESARSNEE 265
Db 140 AQENHEEVVKYLLKKGANQALSTEDGFTPLAVA-----LQOQHDR 179
QY 266 KMMALLPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGLVPLHN 324
Db 180 VVAVLLE-----NDSKGVRLPALHIAAKDDTTAATLLQNEHNPDTVTSKSGFTPLHI 233
QY 325 ACSGYHYEVTLLVKHGACVNMALWQFTPLHEAAKSNRVEVCSLLSYGADPTLLNCHN 384
Db 234 AAHYGHENVQQLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IIDSRT 290
QY 385 KSAIDLATTPOLKERLAVFEFGHLSLLQAREADVTRIKKHSLEVMVNFKHQPTHETALHC 444
Db 291 KDL-----TPLHC 299
QY 445 AAASPYPRKQTCIELLRKGANINEKTEFLTPLHVASEKAHNDVVVVKHEAKVN--A 502
Db 300 AARSGH---DQVOLLVVGQAPISAKTKNGLAPLHMAAGDHVDAARTLLYHRAVDVDT 356
QY 503 LDNLGQTSLHRAAYCGHLOTCRLLLSYGCDPNIISLQGFALOMNGEN----- 550
Db 357 VDYL--TPLHVAACHGVRAKLLLDLDRSDPNRSALNGFTPLHIACKKNRKIVKVELLYK 414

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 21:15:54 ; Search time 15.3675 Seconds
(without alignments)
3346.711 Million cell updates/sec

Title: US-09-843-159B-4

Perfect score: 6464

Sequence: 1 RCSARRGAGGQAGRGARY.....AYPEYLITYQIMRPEGMVDG 1240

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6108	94.5	1166	1	TNK2_HUMAN
2	5103.5	79.0	1327	1	TNK1_HUMAN
3	892.5	13.8	4377	1	ANK3_HUMAN
4	878.5	13.6	1880	1	ANK1_HUMAN
5	877	13.6	3924	1	ANK2_HUMAN
6	874	13.5	1862	1	ANK1_MOUSE
7	760.5	11.8	1059	1	Y379_HUMAN
8	571	8.8	768	1	YB23_HUMAN
9	485	7.5	1401	1	LATA_LATMA
10	431	6.7	603	1	V162_FOWPV
11	399.5	6.2	747	1	V222_FOWPV
12	390.5	6.0	832	1	ANK3_HUMAN
13	382.5	5.9	596	1	V024_FOWPV
14	380.5	5.9	668	1	V244_FOWPV
15	357	5.5	587	1	ANK2_HUMAN
16	354.5	5.5	776	1	ANK5_HUMAN
17	354.5	5.5	1431	1	DAPK_HUMAN
18	351	5.4	542	1	V155_FOWPV
19	351	5.4	775	1	ANK5_MOUSE
20	327.5	5.1	525	1	V228_FOWPV
21	322	5.0	525	1	ASB3_MOUSE
22	321	5.0	436	1	V245_FOWPV
23	318	4.9	592	1	V246_FOWPV
24	315.5	4.9	656	1	FEMI_CAEEL
25	312.5	4.8	568	1	PPGB_MOUSE
26	310	4.8	528	1	PPGA_HUMAN
27	309	4.8	524	1	PPGA_MOUSE
28	308	4.8	518	1	ASB3_HUMAN
29	307.5	4.8	568	1	PPGB_BOVIN
30	305	4.7	735	1	RN5A_MOUSE
31	304	4.7	567	1	PPGB_HUMAN
32	303	4.7	429	1	AS10_HUMAN
33	300.5	4.6	692	1	ANK6_HUMAN

RESULT 1
TNK2_HUMAN

ID	TNK2_HUMAN	STANDARD:	PRT: 1166 AA.
AC	Q9H2K2; Q9HAS4; Q9H8F2;		
DT	15-JUN-2002 (Rel. 41, Created)		
DT	15-JUN-2002 (Rel. 41, Last sequence update)		
DE	Tankyrase 2 (EC 2.4.2.30) (TANK2) (Tankyrase II) (TNKS-2) (TRF1-interacting ankyrin-related App-ribose polymerase 2) (Tankyrase-like protein) (Tankyrase-related protein).		
DE	TNK52 OR TNKL OR TANK2.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_taxid=9606;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	TISSUE=Fetal brain;		
RX	MEDLINE=21072550; PubMed=11205898;		
RA	Monz D., Munnia A., Comtesse N., Fischer U., Steudel W.-I., Feiden W., Glass B., Meese E.U.;		
RA	"Novel tankyrase-related gene detected with meningioma-specific sera.";		
RT	Clin. Cancer Res. 7:113-119(2001).		
RL	[2]		
RC	SEQUENCE FROM N.A.		
RC	TISSUE=Breast Carcinoma;		
RX	MEDLINE=21190090; PubMed=11294570;		
RA	Kulmov A.N., Kuprash D.V., Petrov V.N., Vdovichenko K.K., Scanlan M.J., Jongeneel C.V., Lagarkova M.A., Nedospasov S.A.;		
RT	"Cloning and characterization of TNKL, a member of tankyrase gene family.";		
RL	Genes Immun. 2:52-55(2001).		
RN	[3]		
RC	SEQUENCE FROM N.A., AND SUBCELLULAR LOCALIZATION.		
RC	TISSUE=Liver;		
RX	MEDLINE=21264473; PubMed=11278563;		
RA	Lyons R.J., Deane R., Lynch D.K., Ye Z.-S.J., Sanderson G.M., Eyre H.J., Sutherland G.R., Daly R.J.;		
RT	"Identification of a novel human tankyrase through its interaction with the adaptor protein Grb14.";		
RL	J. Biol. Chem. 276:17172-17180(2001).		
RN	[4]		
RC	SEQUENCE FROM N.A., AND SUBCELLULAR LOCALIZATION.		
RC	TISSUE=Placenta;		
RX	MEDLINE=21443728; PubMed=11454873;		
RA	Kaminker P.G., Kim S.-H., Taylor R.D., Zebarjadian Y., Funk W.D., Morin G.B., Yaswen P., Campisi J.;		
RT	"TANK2, a new TRF1-associated poly(Adp-ribose) polymerase, causes rapid induction of cell death upon overexpression.";		
RL	J. Biol. Chem. 276:35891-35899(2001).		
RN	[5]		
RC	SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH TRF1 AND LNPEP/OTASE.		
RC	TISSUE=Skeletal muscle;		
RX	MEDLINE=21661461; PubMed=11802774;		

34	291.5	4.5	741	1	RN5A_HUMAN	Q05823	homo sapien
35	291.5	4.5	950	1	ORP1_HUMAN	Q9BXW6	homo sapien
36	285	4.4	461	1	V218_FOWPV	Q9J517	fowlpox vir
37	285	4.4	1083	1	Y1L2_YEAST	P40480	saccharomyc
38	283.5	4.4	226	1	PSDA_HUMAN	Q75832	homo sapien
39	281.5	4.4	434	1	V219_FOWPV	Q9J516	fowlpox vir
40	281.5	4.4	482	1	V232_FOWPV	Q9J503	fowlpox vir
41	281.5	4.4	898	1	KBF2_HUMAN	Q00653	homo sapien
42	280.5	4.3	231	1	PSDA_RAT	Q922X3	rattus norv
43	278	4.3	984	1	KBF1_CHICK	Q04861	gallus gall
44	277	4.3	1964	1	NTC4_MOUSE	P31695	mus musculu
45	275	4.3	347	1	GABC_MOUSE	Q00421	mus musculu

ALIGNMENTS

Db 481 LOEGISLGNSEADROLLAANKAGDVETVKLCTVGSVNCRDIEGRQSTPLHFAAGYNRV5 540
 QY 615 VVEYLLOHGADVHAKDKGLPLHNACSYGHVEAELLVKGAVVNVADLWKFTPLHEAA 674
 Db 541 VVEYLLOHGADVHAKDKGLPLHNACSYGHVEAELLVKGAVVNVADLWKFTPLHEAA 600
 QY 675 AKGYEICKLLOHGADPTKKNRQDNTPLDLVKDGDITDIDLLRGAALLDAAKKGCLAR 734
 Db 601 AKGYEICKLLOHGADPTKKNRQDNTPLDLVKDGDITDIDLLRGAALLDAAKKGCLAR 660
 QY 735 VKLSSPDVNCRDQGRHSTPLHLAAGYNVLEVAEYLLOHGADVNAQDKGLPLHNA 794
 Db 661 VKLSSPDVNCRDQGRHSTPLHLAAGYNVLEVAEYLLOHGADVNAQDKGLPLHNA 720
 QY 795 SYGHVDVAALLIKYNACYNATDKWAFTPLHEAAQKGRTOCALILLAHGADPTLKNOEGQT 854
 Db 721 SYGHVDVAALLIKYNACYNATDKWAFTPLHEAAQKGRTOCALILLAHGADPTLKNOEGQT 780
 QY 855 PLDLVSADVDVALLTAAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 914
 Db 781 PLDLVSADVDVALLTAAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 840
 QY 915 LDNLSSGSELSVVSSSGTEGASLEKKEVPGVDFTSIQFVRNLGLBHLMDIFEREQIT 974
 Db 841 LDNLSSGSELSVVSSSGTEGASLEKKEVPGVDFTSIQFVRNLGLBHLMDIFEREQIT 900
 QY 975 LDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPYLTLNTSGSGTILIDLSP 1034
 Db 901 LDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPYLTLNTSGSGTILIDLSP 960
 QY 1035 DDKFQSVSEEMQSTVREHRDGGHAGGTFFNRYNLIKQKCNKKLWERYTHRRKEVSEEN 1094
 Db 961 DDKFQSVSEEMQSTVREHRDGGHAGGTFFNRYNLIKQKCNKKLWERYTHRRKEVSEEN 1020
 QY 1095 HNANERMLFHGSPFNVAIHKGFDERHAYIGGMFGAGIYFAENSSKNQVYIGGGTG 1154
 Db 1021 HNANERMLFHGSPFNVAIHKGFDERHAYIGGMFGAGIYFAENSSKNQVYIGGGTG 1080
 QY 1155 CPVHKDRSCYICHRQLLCFRVTLGKSFQFSAMKMAHSPGPHSHSVTGRPSVNGLALAEYV 1214
 Db 1081 CPVHKDRSCYICHRQLLCFRVTLGKSFQFSAMKMAHSPGPHSHSVTGRPSVNGLALAEYV 1140
 QY 1215 IYRGEQAYPELYITQIMRPGMVDG 1240
 Db 1141 IYRGEQAYPELYITQIMRPGMVDG 1166

RESULT 2
 TNK1_HUMAN STANDARD; PRT; 1327 AA.
 ID TNK1_HUMAN
 AC O95271; O95272;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Tankyrase 1 (EC 2.4.2.30) (TANK1) (tankyrase 1) (TNKS-1) (TRF1-
 DE interacting ankyrin-related ADP-ribose polymerase).
 GN TNKS OR TNKSI OR TIN1 OR TINFI OR PARPL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Testis;
 RX MEDLINE=99040105; PubMed=9822378;
 RA Smith S., Giriat I., Schmitt A., de Lange T.:
 RT "Tankyrase, a poly(ADP-ribose) polymerase at human telomeres.";
 RL Science 282:1484-1487 (1998).
 RN [2]
 RP SUBCELLULAR LOCALIZATION.
 RX MEDLINE=99454782; PubMed=10523501;
 RA Smith S., de Lange T.:
 RT "Cell cycle dependent localization of the telomeric PARP, tankyrase,

to nuclear pore complexes and centrosomes.";
 J. Cell Sci. 112:3649-3656 (1999).
 [3]
 FUNCTION, AND PHOSPHORYLATION.
 MEDLINE=20556282; PubMed=10988299;
 Chi N.-W., Lodish H.F.:
 RT "Tankyrase is a golgi-associated mitogen-activated protein kinase
 RT substrate that interacts with IRAP in GLUT4 vesicles.";
 J. Biol. Chem. 275:38437-38444 (2000).
 [4]
 FUNCTION, AND MUTAGENESIS OF HIS-1184 AND GLU-1291.
 MEDLINE=21602874; PubMed=11739745;
 Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.:
 RT "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2
 RT at human telomeres.";
 Mol. Cell. Biol. 22:332-342 (2002).
 CC -!- FUNCTION: May regulate vesicle trafficking and modulate the
 CC subcellular distribution of SLC24A4/GLUT4-vesicles. Has PARP
 CC activity and can modify TRF1, and thereby contribute to the
 CC regulation of telomere length.
 CC -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribose)1(N)-acceptor =
 CC nicotinamide + (ADP-D-ribose)1(N+1)-acceptor.
 CC -!- SUBUNIT: Oligomerizes and associates with TNKS2. Interacts with
 CC the cytoplasmic domain of LNPP/otase in SLC24A4/GLUT4-vesicles.
 CC Binds to the N-terminus of telomeric TRF1 via the ANK repeats.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
 CC with juxtanuclear SLC24A4/GLUT4-vesicles. A minor proportion is
 CC also found at nuclear pore complexes and around the pericentriolar
 CC matrix of mitotic centrosomes. During interphase, a small fraction
 CC of TNKS is found in the nucleus, associated with TRF1.
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
 CC produced by alternative splicing.
 CC -!- TISSUE SPECIFICITY: Ubiquitous; highest levels in testis.
 CC -!- PTM: Upon insulin-stimulation, phosphorylated on serine residues
 CC by MAPK kinases.
 CC -!- PTM: ADP-ribosylated (-auto).
 CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
 CC -!- SIMILARITY: CONTAINS 15 ANK REPEATS.
 CC -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; AF082556; AAC79841.1; -
 DR EMBL; AF082557; AAC79842.1; -
 DR EMBL; AF082558; AAC79843.1; -
 DR EMBL; AF082559; AAC79844.1; -
 DR HSSP; Q00420; IAWC.
 DR Genew; HGNC:11941; TNKS.
 DR MIM; 603303; -
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR001660; SAM.
 DR Pfam; PF00023; ank; 21.
 DR Pfam; PF00536; SAM; 1.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 15.
 DR SMART; SM00454; SAM; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 15.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PS50105; SAM_DOMAIN; 1.
 DR Transferrase; Glycosyltransferase; NAD; Golgi stack; Telomere;
 DR Nuclear protein; Repeat; ANK repeat; ADP-ribosylation;
 DR Phosphorylation; Alternative splicing.
 FT REPEAT 215 247 ANK 1.
 FT REPEAT 248 280 ANK 2.
 FT REPEAT 281 313 ANK 3.
 FT REPEAT 368 400 ANK 4.
 FT REPEAT 401 433 ANK 5.

```

FT REPEAT 434 466 ANK 6.
FT REPEAT 521 556 ANK 7.
FT REPEAT 557 589 ANK 8.
FT REPEAT 590 622 ANK 9.
FT REPEAT 683 715 ANK 10.
FT REPEAT 716 748 ANK 11.
FT REPEAT 749 781 ANK 12.
FT REPEAT 836 868 ANK 13.
FT REPEAT 869 901 ANK 14.
FT REPEAT 902 934 ANK 15.
FT DOMAIN 1030 1089 SAM.
FT DOMAIN 1176 1327 PARP.
FT DOMAIN 9 14 POLY-HIS.
FT DOMAIN 27 34 POLY-PRO.
FT DOMAIN 128 134 POLY-SER.
FT DOMAIN 137 145 POLY-SER.
FT VARSPIC 641 643 EST -> GHS (IN ISOFORM 2).
FT VARSPIC 644 1327 MISSING (IN ISOFORM 2).
FT MUTAGEN 1184 1184 H->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
FT MUTAGEN 1291 1291 E->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
FT MUTAGEN 1291 1291 WITH A-1184.
SQ SEQUENCE 1327 AA; 142010 MW; 5140985C710B957 CRC64;

Query Match 79.0%; Score 5103.5; DB 1; Length 1327;
Best Local Similarity 79.2%; Pred. No. 1.6e-299;
Matches 963; Conservative 110; Mismatches 128; Indels 15; Gaps 3;

QY 22 AAGTAPDVTAGSQAARALSASSPGGLALLAGPGLLRLLRLLALLAVAAARTMSGRCA 81
DB 112 SAAGVAPNPAGSGNNSSPSSTSSSSSSPG-----SSLAESPEAACVSTAPL 165
QY 82 GGAACASAAAEEAVEPAARELFACRNGDVERVYKRLVTPPEKYNSTRDTAGKSTPLHFAAG 141
DB 166 GPCAGAGTGVPAGVAGALRELEACRNGDVSRYKRLVDAANVAKMDACKRKSPLHFAAG 225
QY 142 FGRKVVYELLQGANVQARDGGLIPLHNACSFHGAEEVNVLLRHGADPNARDNNYTP 201
DB 226 FGRKVVYELLQGANVHARDGGLIPLHNACSFHGAEEVNVLLRCOGADPNARDNNYTP 285
QY 202 LHEAAIKGIDVICVLQGAEPTRTGTALDADPSAKAVLTGEYKKDELLESARS 261
DB 286 LHEAAIKGIDVICVLQGAEPTRTGTALDADPSAKAVLTGEYKKDELLESARS 345
QY 262 GNEEKWMLLTPLVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAOKDGLVP 321
DB 346 GNEEKWMLLTPLVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAOKDGLVP 405
QY 322 LHNACSYGHYEVTELLVKKHCACVANDLQWFTPLHEAASKNRVEVCSLLLSYGADPTLN 381
DB 406 LHNACSYGHYEVTELLVKKHCACVANDLQWFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465
QY 382 CHNKAIDLAPTQPKERLAYEPKSHLSLQAAREADVTRIKKHLSEVMVNFKHPQTHETA 441
DB 466 CHKSASVDMAPTPELRLETYEPKSHLSLQAAREADLAKVKTLEIINFKQPSHETA 525
QY 442 LHCAASAPPKRQIOICELLRLKGANINEKTKPELTPLHVASEKAHNDVVEVVKHBAKVN 501
DB 526 LHCAVASLHPKRRQVTELLRLKGANINEKTKPELTPLHVAAEAHNDVMEVLKHKGAKMN 585
QY 502 ALDNLGQTSILHRAAYGHLQTCRLLLSYGCDDPNTISLOGFTALQMGNNVQQLLQEGISL 561
DB 586 ALDTLGTQALHRAALAGHLQTCRLLLSYGSDDPSIISLQGTAAQMGNEAVQQLLSSTPI 645
QY 562 GNSEADRLLEAAKAGDVETVKKIKCTVQSVNCRDIEGROSTPLHFAAGYNRVSVWEYLLQ 621
DB 646 RTSVDVYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVWEYLLH 705
QY 622 HGADVHAOKDGLVPLHNACSYGHYEVTELLVKKHGVVNVADLWKTTPPLHEAAAKGYEI 681
DB 706 HGADVHAOKDGLVPLHNACSYGHYEVTELLVKKHGVVNVADLWKTTPPLHEAAAKGYEI 765
QY 682 CKLLQHGADPTKKNRDGNTPLDLVRDGDPTDIQDLLRGDAALLDAACKGCLARVKKLSSP 741

```

```

DB 766 CKLLKLGADPTKKNRDGNTPLDLVRKEGDPTDIQDLLRGDAALLDAACKGCLARVKKLCTP 825
QY 742 DNVNCRDPTGHRSTPLHLAAGYNVNLLEVAEYLLQHGADVNAQDKGGLIPLHNAAASYGHVDV 801
DB 826 ENINCRDPTGHRSTPLHLAAGYNVNLLEVAEYLLQHGADVNAQDKGGLIPLHNAAASYGHVDI 885
QY 802 AALLIKYNACVNATDKWAFTPLHEAAQKRTQLCALLAHGADPTPLKNOEGOTPLDLVSA 861
DB 886 AALLIKYNACVNATDKWAFTPLHEAAQKRTQLCALLAHGADPTPLKNOEGOTPLDLVSA 945
QY 862 DVSALLTAAMPSPALPSYCKQVNLGVRSPPGATADALSSGSSPSSLSAASLDNLSSG 921
DB 946 DDIRALLIDAMPPEALPTCFKQAT-----VVSASLISPASTPSCLSAASIDNLTPG 998
QY 922 FSELSVSVSSGTEGASLSLEKK--EVPGVDFSTQVRNLGLEHLMDFEREQITLDDVLV 979
DB 998 LAELAVGGASNAGDGAAGTERGEVAGLDNMISQFLKSLGLEHLDIPETQITLDDVLA 1058
QY 980 EMGHKELKEIGINAYGHRHKLKINGVERLISGQOGLNPYLNTSGSGTILIDLSPPDKKEF 1039
DB 1059 DMGHEELKEIGINAYGHRHKLKINGVERLISGQOGLNPYLNTSGSGTILIDLSPPDKKEF 1118
QY 1040 QSVVEEMQSTVREHRDGGHAGGIFNRYNLIKQVCNKKLWERYTHRRKEVSEENHNAN 1099
DB 1119 QSVVEEMQSTVREHRDGGHAGGIFNRYNLIKQVCNKKLWERYTHRRKEVSEENHNAN 1178
QY 1100 ERMFLHGSFPVNAIHKGFDERHAYIGGMFGAGIYFAENSCKSNQYVYIGGGTGCPVHK 1159
DB 1179 ERMFLHGSFPVNAIHKGFDERHAYIGGMFGAGIYFAENSCKSNQYVYIGGGTGCPVHK 1238
QY 1160 DRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPOHSHSVTGRPSVNGALARYIYRGE 1219
DB 1239 DRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPOHSHSVTGRPSVNGALARYIYRGE 1298
QY 1220 QAYPEVLTITQIMRPE 1235
DB 1299 QAYPEVLTITQIMRPE 1314

RESULT 3
ANK3_HUMAN
ID ANK3_HUMAN STANDARD; PRT; 4377 AA.
AC Q12955;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 3 (ANK-3) (Ankyrin G).
GN ANK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain stem;
RX MEDLINE=95138209; PubMed=7836469;
RA Kordeli E., Lambert S., Bennett V.;
RT "Ankyrin G. A new ankyrin gene with neural-specific isoforms localized
RT at the axonal initial segment and node of Ranvier.";
RL J. Biol. Chem. 270:2352-2359(1995).
CC -!- FUNCTION: Membrane-cytoskeleton linker.
CC -!- ALTERNATIVE PRODUCTS: A number of isoforms are produced by
CC alternative splicing.
CC -!- TISSUE SPECIFICITY: Expressed in brain and other tissues.
CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```


RN [3]
 RP VARIANT HS ILF-462.
 RX MEDLINE-96225450; PubMed-8640229;
 RA Eber S.W.; Gonzalez J.M.; Lux M.L.; Scarpa A.L.; Tse W.T.,
 RA Dornwell M.; Herbers J.; Kugler W.; Oezcan R.; Pekrun A.,
 RA Gallagher P.G.; Schroeter W.; Forget B.G.; Lux S.E.;
 RT "Ankyrin-1 mutations are a major cause of dominant and recessive
 RT hereditary spherocytosis.";
 RL Nat. Genet. 13:214-218(1996).
 CC -!- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
 CC ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
 CC NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE
 CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
 CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
 CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
 CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
 CC PLASMA MEMBRANE.
 CC -!- ALTERNATIVE PRODUCTS: At least 3 isoforms; 1/2.1 (shown here),
 CC 2/2.2 and 3; are produced by alternative splicing.
 CC -!- PTM: REGULATED BY PHOSPHORYLATION.
 CC -!- PTM: ACYLATED BY PALMITIC ACID GROUP(S).
 CC -!- DISEASE: Defects in ANK1 are the cause of dominant and recessive
 CC hereditary spherocytosis (HS).
 CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.
 CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X16609; CAA34610.1; -;
 DR EMBL; M28880; AAA51732.1; -;
 DR PIR; S08275; SJHUK.
 DR PIR; A35049; A35049.
 DR HSSP; Q00420; IAWC.
 DR Genew; HGNC:492; ANK1.
 DR MIM; 182900; -;
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00023; ank; 24.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 22.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 20.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
 KW Phosphorylation; Lipoprotein; Disease mutation; Elliptocytosis;
 KW Polymorphism.
 FT INIT_MET 0 0
 FT DOMAIN 1 826 89 KDA DOMAIN (ANION EXCHANGE PROTEIN
 FT BINDING DOMAIN).
 FT DOMAIN 827 1381 62 KDA DOMAIN (SPECTRIN BINDING
 FT DOMAIN).
 FT DOMAIN 1382 1880 55 KDA REGULATORY DOMAIN (REGULATES
 FT THE BINDING OF ANKYRIN TO SPECTRIN
 FT AND THE BAND 3 PROTEIN).
 FT REPEAT 43 72 ANK 1.
 FT REPEAT 76 105 ANK 2.
 FT REPEAT 109 138 ANK 3.
 FT REPEAT 142 171 ANK 4.
 FT REPEAT 173 200 ANK 5.
 FT REPEAT 204 233 ANK 6.
 FT REPEAT 237 266 ANK 7.

FT REPEAT 270 299 ANK 8.
 FT REPEAT 303 332 ANK 9.
 FT REPEAT 336 365 ANK 10.
 FT REPEAT 369 398 ANK 11.
 FT REPEAT 402 431 ANK 12.
 FT REPEAT 435 464 ANK 13.
 FT REPEAT 468 497 ANK 14.
 FT REPEAT 501 530 ANK 15.
 FT REPEAT 534 563 ANK 16.
 FT REPEAT 567 596 ANK 17.
 FT REPEAT 600 629 ANK 18.
 FT REPEAT 633 662 ANK 19.
 FT REPEAT 666 695 ANK 20.
 FT REPEAT 699 728 ANK 21.
 FT REPEAT 732 761 ANK 22.
 FT REPEAT 765 794 ANK 23.
 FT DOMAIN 1402 1486 DEATH.
 FT VARSPLIC 1512 1873 MISSING (IN ISOFORM 2).
 FT VARSPLIC 1874 1880 H -> D (IN ISOFORM 2).
 FT VARSPLIC 1849 1880 TVEGPLEDSELEVDIDYFMKSHKDHSTPNP -> ELRGS
 FT 3). GLOPLIEGRKGAQIVKRASLRKRGK (IN ISOFORM
 FT 3).
 FT R -> T.
 FT /FTId=VAR_000595.
 FT V -> I (IN HS).
 FT /FTId=VAR_000596.
 FT R -> H (IN BRUEGGEN).
 FT /FTId=VAR_000597.
 FT V -> A.
 FT /FTId=VAR_000598.
 FT D -> E.
 FT /FTId=VAR_000599.
 FT E -> D.
 FT /FTId=VAR_000601.
 FT S -> T.
 FT /FTId=VAR_000600.
 FT D -> N (IN DUESSELDORF).
 FT /FTId=VAR_000602.
 FT R -> D.
 FT /FTId=VAR_000603.
 FT A -> S (IN REF. 2).
 FT CONFLICT 229 229
 FT CONFLICT 1545 1545 V -> I (IN REF. 2).
 FT SEQUENCE 1880 AA; 206145 MW; 1C5F5E7EFDICD428 CRC64;
 SQ
 Query Match 13.6%; Score 878.5; DB 1; Length 1880;
 Best Local Similarity 25.8%; Pred. No. 8.2e-45;
 Matches 292; Conservative 157; Mismatches 375; Indels 307; Gaps 28;
 QY 105 ACRNGDVERVRLVTPKVNSTRDTAGKSTPLHFAAGFGKRDVVEYLLONGANYQARDG 164
 Db 51 ASKEGHVKMVELLHKE-IIETTTKKGNTALHIAALAGODEVVRVRELNYGANYNAOSQK 109
 QY 165 GLIPLHNACSFHAEVYVNVLLLRHCGADPNARDNNWYTPLHEAAIKGKIDVCIVLQHAEP 224
 Db 110 GTPLXMAAENHLEVVRKFLLENGANONVATEDGFTPLAVALQOGHENVVAHLINYGTKG 169
 QY 225 TI-----RNTDGRALTDLADPSAKAVITGEYKKDELIESARSNGNEEKMMALLPLN 275
 Db 170 KVRLPALHIAARNDDTRTA-----AVLLQNDPNPDL--SKTG-----FTPLH 210
 QY 276 VNCIASGRKS-----TPHLAAGYNRVKIVOLLLOHGDVHAKDKG 317
 Db 211 IAAHYENLVNVAQLLLNRGASVNFPTQGITPLHIASRRGNVIMVRLLDORGAQTETKTD 270
 QY 318 DLVPLHNACSGHYEVTELVKHCACVNDLWQFTPLHEAASKNRVVCVSLLSLYGADP 377
 Db 271 ELTEPLHCAARNGHVRISSEILDHGAIPQAKTKNGLSPHMAAOGDHLDCVRLLLQYDAE- 329
 QY 378 TLLNCHNKSATDLAPTQPKERLAYEFKGHSLLOAAREADVTRIKKHLSEMVNFKHPQT 437
 Db 330 -----ID-----DIT-----LD 336
 QY 438 HETALHCAAAAPYPRKQICELLRLKGANINEKTKFTPLHVASEKAHNDVVEVVKHE 497

```

Db 337 HLTPLHVA--HCGHRRYAKVLLDKGAPNSGRALNGFTPLHIAACKNHRVVMELLLKKG 393
Qy 498 AKVNALDNLGOTSLHRAAYCGHLQTCRLLSYGCOPNILLSQOFTALOMG-----NENYQ 552
Db 394 ASIDAVTESGLTPLHVASPMGHPVIVKLLQORQAGPNVSNVKTPLHMAARAGHTEVAK 453
Qy 553 QLLQE-----GISLGNSEADRLLE-----AAKAG 577
Db 454 YLLQKAKYNAKAKDDQTPHCAARIGHTNMVKLLLENNANPNLATTAGTTPHIAARSG 513
Qy 578 DYETVKKICTVO-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDGGLVP 636
Db 514 HVETVLALLEKESAOACMTKG--FTPLHVAAYKGVRAEELLERDAHPNAAGKNGLTP 571
Qy 637 LHNACSYGHEVAEYLLHGVAVNVADLWK-FTPLHEAAAKGYEICKLLQHGADPTTK 695
Db 572 LHVAVHNNLDIVKLLPRGSGPH-SPAWNGTTPHIAAKQVQVARSLLQYGSANAE 630
Qy 696 NRDGNTPDL-LVKDGTDTIDQLLGDAAALLDAKGLARVKLSLSPDNVNCRTDQGRHS 754
Db 631 SVQGVTPHLAAQEGHAEMVALL-----LSKOANGNLGKSG--L 668
Qy 755 TPLHLAGAGNNLEVAEYLLQHGADYNAQDKGLIPLHNAASGHVDVAALLIKYNACVNA 814
Db 669 TPLHLVAQEGHVPADVLLKHGMVMDATRMGYTPLHVASHYGNIKLVKFLQHOADVNA 728
Qy 815 TDKWFTPLHEAAQKRGRTOLCALLAHGADPTLKNOEGOTPLDL-----VS 860
Db 729 KTKLGYSPHQAQOOGHTDVTLLKNGASPNVSDGTTPLAIKRLGYISVTDVLKVV 788
Qy 861 ADVSALLTAAMPSPALPSC-----YKPO 884
Db 789 TDETSFVLVSDKHRMSFPETVDEILDVSEGEELISFAERRDSRDVDEKELLDVFPK 848
Qy 885 VLVNGVRSFCATADALSSGSPSSLS-----AASLDNLGSFSELSVSVSSGTEGASS 939
Db 849 LDQVVESPAII-----PRIPCAMPTVIRSEEQEAKSEYDEDSLIPSPATETSDN 900
Qy 940 LEKKVQPVDFSIITOFVNLGLEHLMDFEREQITLDVLVEMGHKELKEIGNAYGHRHK 999
Db 901 ISPVASPVHTGFLVSPM-----VDRGSGMRGSRHN 931
Qy 1000 LIKVE-----RLISGQOGLNPVYLTNTSGTTLIDLSP-----1034
Db 932 GLRVVIPPRTCAAPTRITCRVKPKLSTPPLPABEEGLASRIIALGPTGAQFLSPVIVE 991
Qy 1035 -----DKKEFOSVEEMOSTVREHRCGHAGGIENRYNLIKIQVCN 1076
Db 992 IPHFASHGRGDRELVLVLRSENGSVWKEHR-----SRYGESYLDQILN 1033

RESULT 5
ANK2_HUMAN
ID ANK2_HUMAN STANDARD; PRT: 3924 AA.
AC Q01484; Q01485;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1; 2 AND 3).
RC TISSUE=Brain stem;
RX MEDLINE=91302466; PubMed=1830053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain
RT ankyrins reveal a family of alternatively spliced genes.";
RL J. Cell Biol. 114:241-253(1991).

[2]
RP REVISIONS.
RA Carpenter S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain stem;
RX MEDLINE=94075409; PubMed=8253844;
RA Chan W., Kordeli E., Bennett V.;
RT "440-kD ankyrinB: structure of the major developmentally regulated
RT domain and selective localization in unmyelinated axons.";
RL J. Cell Biol. 123:1463-1473(1993).
[4]
RP SEQUENCE OF 463-495 FROM N.A.
RX MEDLINE=92009921; PubMed=1833308;
RA Tse W.T., Menninger J.C., Yang-Peng T.L., Francke U., Sahr K.E.,
RA Lux S.E., Ward D.C., Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid
RT ankyrin gene.";
RL Genomics 10:858-866(1991).
CC -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC elements. Also bind to cytoskeletal proteins.
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2 AND 3; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL
CC CELLS THROUGHOUT THE BRAIN.
CC -!- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES
CC AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
CC AND FUNCTION (POTENTIAL).
CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL: X56957; CAA40278.1; -
CC DR EMBL: X56958; CAA40279.2; -
CC DR EMBL: Z26634; CAB42644.1; -
CC DR EMBL: M37123; AAA62828.1; -
CC DR PIR: S14533; S14533.
CC DR PIR: A39643; A39643.
CC DR PIR: B39643; B39643.
CC DR HSSP: S14569; S14569.
CC DR HSSP: P42771; IDC2.
CC DR Genew; HGNC:493; ANK2.
CC DR MIM: 106410; -
CC DR InterPro; IPR002110; ANK.
CC DR InterPro; IPR000488; Death.
CC DR InterPro; IPR000906; ZUS.
CC DR Pfam; PF00023; ank; 24.
CC DR Pfam; PF00531; death; 1.
CC DR Pfam; PF00791; ZUS; 1.
CC DR PRINTS; PR01415; ANKYRIN.
CC DR SMART; SM00248; ANK; 21.
CC DR SMART; SM00005; DEATH; 1.
CC DR SMART; SM00218; ZUS; 1.
CC DR PROSITE; PS50088; ANK_REPEAT; 20.
CC DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC DR PROSITE; PS50017; DEATH_DOMAIN; 1.
CC DR Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
CC Phosphorylation.
CC REPEAT 63 92 ANK 1.
CC FT REPEAT 96 125 ANK 2.
CC FT REPEAT 129 158 ANK 3.
CC FT REPEAT 162 191 ANK 4.
CC FT REPEAT 193 220 ANK 5.
CC FT REPEAT 232 261 ANK 6.
CC FT REPEAT 265 294 ANK 7.

```

```

FT REPEAT 298 327 ANK 8.
FT REPEAT 331 360 ANK 9.
FT REPEAT 364 393 ANK 10.
FT REPEAT 397 426 ANK 11.
FT REPEAT 430 459 ANK 12.
FT REPEAT 463 492 ANK 13.
FT REPEAT 496 525 ANK 14.
FT REPEAT 529 558 ANK 15.
FT REPEAT 562 591 ANK 16.
FT REPEAT 595 624 ANK 17.
FT REPEAT 628 657 ANK 18.
FT REPEAT 661 690 ANK 19.
FT REPEAT 694 723 ANK 20.
FT REPEAT 727 756 ANK 21.
FT REPEAT 760 789 ANK 22.
FT REPEAT 793 822 ANK 23.
FT DOMAIN 1773 1950 REPEAT-RICH REGION.
FT REPEAT 1773 1784 REPEAT A.
FT REPEAT 1785 1796 REPEAT A.
FT REPEAT 1797 1808 REPEAT A.
FT REPEAT 1809 1820 REPEAT A.
FT REPEAT 1821 1832 REPEAT A.
FT REPEAT 1833 1844 REPEAT A.
FT REPEAT 1845 1856 REPEAT A.
FT REPEAT 1857 1867 REPEAT A.
FT REPEAT 1868 1879 REPEAT A.
FT REPEAT 1880 1891 REPEAT A.
FT REPEAT 1892 1902 REPEAT A.
FT REPEAT 1903 1914 REPEAT A.
FT REPEAT 1915 1926 REPEAT A.
FT REPEAT 1927 1938 REPEAT A.
FT REPEAT 1939 1950 REPEAT A.
FT DOMAIN 3536 3620 DEATH.
FT VARSPLIC 1039 1039 Q -> QFLGKLHLPTAPPPLNEGESLVRILQLGPCTK
      (IN ISOFORM 2).
FT VARSPLIC 1444 3528 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT CONFLICT 475 476 GQ -> PE (IN REF. 4).
FT CONFLICT 971 971 I -> S (IN REF. 1).
FT CONFLICT 3581 3582 QY -> HA (IN REF. 1).
FT CONFLICT 3586 3586 I -> Y (IN REF. 1).
FT SEQUENCE 3924 AA; 430337 MW; 52AC496C428E29D2 CRC64;

Query Match 13.6%; Score 877; DB 1; Length 3924;
Best Local Similarity 29.7%; Pred. No. 2.9e-44;
Matches 293; Conservative 137; Mismatches 402; Indels 156; Gaps 23;

QY 105 ACRNGDVER-VKRLVPEKYNRSDTAGKSTPLHFAAGFGKRDVVEYLLONGANVQARDD 163
DB 38 AARAGNLDKVVEYLGKGDIDINTCNGLNA--LHLAAKEGHVGLVQELLGRGSSVDSATK 95
QY 164 GGLIPLHNACSFHAEVNNLLRHGADPNARNNNVTPLEHAAIKGKIDYICVILLOHCAE 223
DB 96 KGNTALHIALAQAEVVKVVLVEGANINAQSONGFTPLVMAQENHIDVVKYLLENGAN 155
QY 224 PTIRNTDGR-----ALDLADPSAKAVLTGYKKDEL-----LESARSGNERKMMALLTPLN 275
DB 156 QSTATEDGFTPLAVALQQHQAVALLENDTKGKVRPLPALHIAARKDDTKSAALLIQ-- 213
QY 276 VNCASDGRKS-----TPHLAAGYNRVKIVOLLQHGADVHAKDGLVPLPHN 324
DB 214 -NDHNADVQSKMVMNRTESGFTPLHIAAHYGNVNVATLLNLRGAADVFTARNGITPLHV 272
QY 325 ACSYGHVEVTELLVKHGACVNMALMQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
DB 273 ASKRGNTNWKVLLLDGRGQIDAKTRDGLTPLHCAARSGHQVVELLLERGA-PLLARTKN 331
QY 385 KSAIDLAPTQPKERLAYEKHSLLOQAAREADVTRIKKHLSEWVNFKHP-----QTHE 439
DB 332 -----GLSPLHMAAQGDHVECVKHL-----LQHKAPVDVDTLDYL 366
QY 440 TALHCAASPYPKRKQICETLLLRKGANINEKTEFLTPHVASEKAHNDVVEVVKHEAK 499
DB 367 TALHVA-----HCGHYRVTKLLDKRANPNRNLNGFTPLHIAACKKNRIKMYELLVKYGAS 423

```

```

QY 500 VNALDNLGOTSLHRAAYCGHLOTCRLLLSYQCDPNIISLQGTALOMGN-----ENVOQL 554
DB 424 IQAITESGLTPIHVAAPFMGHLNIVLLLLQNGASPDVTNIRGETALHMAARAGQVEVVRCL 483
QY 555 LQEGI-----SLGNSEADROLLE-----AAKAGDV 579
DB 484 LRNGALVDARAREEQTPHIASRLGKTEIVOLLQOHMAHPDAATNGYTPHLHSAREQV 543
QY 580 ETYVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVVEYLLQHGADVHAKDGLVPLHN 639
DB 544 DVASVLEAGAAHSLATK-KGFTPLHVAAYKGLSDVAKLLQRRRAADSAGKNGLTPLHV 602
QY 640 ACSYGHVEVAELLVKHGAVNVADLMKFTPLHEAAGKYEICKLLOHGHADPTKKNRDG 699
DB 603 AAHYDNQKVALLLLEKASPHATAKNGYTPHIAAKNQMOIASTLLNYGAETNIVTKQG 662
QY 700 NTPDLV-KDGDITDQDLRGDAALDAAKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKW 818
DB 663 VTPHLASQEGHTDMVTLLLDKGANIHMSTKSL-----TSLH 700
QY 759 LAAGYNNLEVAEYLLQHGADVNADQKGLIPLHNAASYGHVDVAALLIKYNACVNATDKW 818
DB 701 LAAQEDKVNVDILTCKHGADODAHTKLGYTPLIVACHYGNVKMVFLLKOGANYNAKTKN 760
QY 819 AFTPLHEAOKGRTQLCALLAHGADPTLKNQEGOTPLDL-----VSADDVSALLTAAM 872
DB 761 GYTPLHQAQOQGHTHIINVLLQHGAKPNATTANGNTALATAKRLGYISVVDTLKVYTEEV 820
QY 873 PPSALPSCYKQVPLNGVRSPGATADALSSGSPSSLSAASSLDNLSPSEL----- 925
DB 821 TTTTITITEK-----HKLNVPTMTTEVL-----DVSDEEGDDTMTGDGGERLPEDLK 868
QY 926 -----SSVSSSGTEGASLEKKEVPGVDFSIQFV--RNLGLEHLMDFEREQITLDVLY 979
DB 869 ELGDDSLPSPQFQLDGMNLYRSLYEGGKSDSLRFSRSHLTLSHAS--YLRDSAVMDDSV 926
QY 980 EMGHKLEKEIGINAYGHRHKLKIGVERL 1007
DB 927 VIPSHQVSTLAKAEARNRYRLSNGTENL 954

RESULT 6
ANK1_MOUSE STANDARD; PRT; 1862 AA.
AC Q02357;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Ankyrin 1 (Erythrocyte ankyrin).
GN ANK1 OR ANK-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Erythrocyte.
RX MEDLINE=92345717; Pubmed=1386265;
RA White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;
RT "Murine erythrocyte ankyrin cDNA: highly conserved regions of the
RT regulatory domain."
RL Mamm. Genome 3:281-285(1992).
CC -!- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
CC ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
CC NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE
CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
CC PLASMA MEMBRANE.
CC -!- PTM: REGULATED BY PHOSPHORYLATION (BY SIMILARITY).

```


RT "Prediction of the coding sequences of unidentified human genes. VII.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RL code for large proteins in vitro.";
 RL DNA Res. 4:141-150(1997).
 CC -1- SIMILARITY: CONTAINS AT LEAST 27 ANK REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB002377; BAA20833.2; -
 DR HSSP: P80144; 2MYO.
 DR InterPro: IPR002110; ANK.
 DR Pfam: PF00023; ank; 28.
 DR SMART: SM00248; ANK; 26.
 DR PROSITE: PS00088; ANK_REPEAT; 24.
 DR PROSITE: PS0297; ANK_REPEAT; 1.
 KW Hypothetical protein; Repeat; ANK repeat.
 FT NON_TER 1
 FT REPEAT 46 75 ANK 1.
 FT REPEAT 79 108 ANK 2.
 FT REPEAT 112 141 ANK 3.
 FT REPEAT 145 174 ANK 4.
 FT REPEAT 178 207 ANK 5.
 FT REPEAT 211 240 ANK 6.
 FT REPEAT 244 273 ANK 7.
 FT REPEAT 277 307 ANK 8.
 FT REPEAT 311 340 ANK 9.
 FT REPEAT 344 373 ANK 10.
 FT REPEAT 377 406 ANK 11.
 FT REPEAT 410 439 ANK 12.
 FT REPEAT 443 472 ANK 13.
 FT REPEAT 476 506 ANK 14.
 FT REPEAT 510 540 ANK 15.
 FT REPEAT 555 584 ANK 16.
 FT REPEAT 588 617 ANK 17.
 FT REPEAT 622 651 ANK 18.
 FT REPEAT 658 687 ANK 19.
 FT REPEAT 691 720 ANK 20.
 FT REPEAT 724 753 ANK 21.
 FT REPEAT 761 790 ANK 22.
 FT REPEAT 793 823 ANK 23.
 FT REPEAT 828 857 ANK 24.
 FT REPEAT 861 891 ANK 25.
 FT REPEAT 895 924 ANK 26.
 FT REPEAT 931 960 ANK 27.
 SQ SEQUENCE 1059 AA; 113465 MW; C1F55E6CFE494770 CRC64;

Query Match 11.8%; Score 760.5; DB 1; Length 1059;
 Best Local Similarity 27.6%; Pred. No. 4.5e-38;
 Matches 275; Conservative 129; Mismatches 357; Indels 235; Gaps 33;
 QY 102 LFEACRNGDVERKRLV-TPEKYNRSRTAGKSTPLHFAAGFGKDVVEYLLONGANVOA 160
 Db 18 LVQAFNGDPDEVALIFPKEDVNFQD--NEKRTPLHAAAYLGDAAEIIILLISGARVNA 75
 QY 161 RDGGGLPLHNACSFHAEVYNLLRHGADPNARD--NWNTPPLHAAIKGIDVCIYLLQ 219
 Db 76 KDSKWLTPHRAVASCSEEAQVQLLHSAVDNARDKNQ--TPLHIAA----- 121
 QY 220 HGAEPTRINDGRTALDAPSAKAVLTGKYEKKDELLESARSNEEKMALLTPLNVNCH 279
 Db 122 -----ANKAV-----KCAEALVPLLSNVN 140
 QY 280 ASDGRKSTPLHAAAYNRKIVOLLQHGADVHAKDKGLVPLHNACSYGHYETELLVK 339
 Db 141 VSDRAGTALHHAFAFSGHGEVKKLLSRGANTINAFDKDRATHAAAYMGHIEVVKLLVS 200
 QY 340 HGACVNAADLWQFTPLHAAAKNRVEYCSLLLSYGADPTLLNCHNKSALDAPTPQLKER 399

Db 201 HGAETVCKDKKSYTPLHAAASSGMISVVKYLLDLGVDNPNAYGNTPHLHVAC----- 253
 QY 400 LAYEFKGHSLLOAAREADVTRIKKHLSL-EMVNFKHPQTHETALHCAAAASPYPKRKOIC- 457
 Db 254 -----YNGQDV-----VNELDCGAIVNOKNEKGF-TPLHFAAASHT---GALCL 295
 QY 458 ELLLRKGANINEKTEFLTPLVHASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYC 517
 Db 296 ELLVGVGADVNMKSKDKTPPLHMTALHGRFSRQTIQSGAVIDCEDKNGTPLHIAARY 355
 QY 518 QHLOTCLRLLSYCCDPNIISLQGFALQMG-----NENVOQLLOEGISLGN-SEADQQL 571
 Db 356 GHELLINTLITSGADTAKRGIHGMFPLHLAALSGFSCCRKLLSSGFDIDTPDFGRTCL 415
 QY 572 EAKA-GDVEYVKKL-----CTVQ-----SYNCRDI 596
 Db 416 HAAAAGNLECLNLLNTGADFNKKDKFGSRPLHYAANANCYQCLFALVSGSASVN--DL 473
 QY 597 EGROSTPLHFAAGYNR-VSVVEYLLQHGADVHAKDKGLVPLHNACSYGHYEAELLVKH 655
 Db 474 DERGCTPLHYAATSDTDGKCLEYLLLRNDANPGIEDRGYNAVHYSAAYGHRCLQLLASE 533
 QY 656 GAV-----VNVADLWK-----FTPLHEAAAKGYEICKLLOLHGADPTKKNRDNTP 703
 Db 534 TPLDVLMTSGTDMLSDSNDRATISPLHLAAYHGHQAELVQLSLDLDRNSSGRTP 593
 QY 704 DLVK-DGDDTDIOLLRGDAALLDAAKKGLARVKKLSSPDNVNCRDPTQGRHSTPLHLAAG 762
 Db 594 DLAAFKGHVECDVVLINQGASI-LVKDYILKR-----TPHAAAT 632
 QY 763 YNNLEVAEYLLQHGADVNA---QDKGLIPLHNAASYGHVDVAALLIKYNACYNATDKWA 819
 Db 633 NGHSECLRLIIGNAEFQNAVYDIQDNGOTPLMLSVLNGHTDCVYSLNKGANVDAKDKWG 692
 QY 820 FTPLHEAAQKGRTOQLCALLAHAGADPTLKNQOGOTPLDLVSA-----DDVSALLTAAMP 875
 Db 693 RTALHRAVGTGHEECVDALLQHGAKCLLRDSRGRTPIHLSAACHIGVLGALLOSAASMD 752
 QY 876 ALP-----SCYKPOVLNG-----VRSPGATADALSSGSPSS 908
 Db 753 ANPATADNHGYTALHWACY-----NGHETVELLLEVEFQKTEGNAFSPHLCAVINDNE 807
 QY 909 LSAASLLDNLSSGSELSVSSSGTEGASL-----EKKEYPGVDFS 951
 Db 808 GAAEMLIDTLG-----ASIVNATDSKGRTPHAAAFTHVECLQLLLSHNAQVNSVD-- 859
 QY 952 ITQFVRNLGLEHLMIDIFEREQI-TLDVLVEMGHKEL 986
 Db 860 -----STGKTPLMMAAENGQTNVEMLVSSASAE 889
 RESULT 8
 YB23_HUMAN
 ID YB23_HUMAN STANDARD; PRT; 768 AA.
 AC Q9ULJ7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein KIAA1223 (Fragment).
 GN KIAA1223.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20039619; PubMed=10574462;
 RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,
 RA Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XV.
 RT The complete sequences of 100 new cDNA clones from brain which code

for large proteins in vitro.*;

DNA Res. 6:337-345(1999).

-!- SIMILARITY: CONTAINS AT LEAST 14 ANK REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AB033049; BAA86537.1; -

DR HSP; P42771; IDC2.

DR InterPro; IPR002110; ANK.

DR Pfam; PF00023; ank; 14.

DR SMART; SM00248; ANK; 13.

DR PROSITE; PS50088; ANK_REPEAT; 13.

DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.

KW Hypothetical protein; Repeat; ANK repeat.

FT NON_TER 1 1

FT REPEAT 1 11 ANK 1.

FT REPEAT 15 44 ANK 2.

FT REPEAT 48 82 ANK 3.

FT REPEAT 86 115 ANK 4.

FT REPEAT 119 148 ANK 5.

FT REPEAT 152 181 ANK 6.

FT REPEAT 185 214 ANK 7.

FT REPEAT 218 247 ANK 8.

FT REPEAT 251 280 ANK 9.

FT REPEAT 284 313 ANK 10.

FT REPEAT 317 346 ANK 11.

FT REPEAT 350 379 ANK 12.

FT REPEAT 383 412 ANK 13.

FT REPEAT 416 446 ANK 14.

SQ SEQUENCE 768 AA; 82819 MW; 2913B69BE2DFE06D CRC64;

Query Match

Best Local Similarity 8.8%; Score 571; DB 1; Length 768;

Matches 203; Conservative 116; Mismatches 262; Indels 256; Gaps 18;

QY 184 LLRHGADPNARDNWNYPPLHEAAIKGIDVCIVILQHGAEPIRNTDGTALDLADPSAK 243

DB 2 LLQHAENVKADNGERTALIAAACHREIVHELDHGAENVHEDVDGRTLSVA----- 56

QY 244 AVLGTGEYKDELLESARSNEEKMALLTPLNVNCHSDGRKSTPLHLAAGYNRVKIVOL 303

DB 57 -----ALCVP-----ASKGHAS-----VVSLL 72

QY 304 LLOHGADYHAKDKGLVPLHNACSYGHYEVTLLYKHGACVYAMDLWQTPPLHEAAKNR 363

DB 73 LIDRGAENVHCDKDGMTPLLVAAAYEGHVDVOLLLEGGADVBDHTDNGRTPLLAASMGH 132

QY 364 VEVCSLLSYGADPTLLNCHNKSADLAPTPOKRLAYEFKGSLLQ-AAREADVTRIK 422

DB 133 ASVWTLFWGA-----AVD-----SIDSGRTVLSTASAGNVEVVR 170

QY 423 KHLSEMVNFKHPOTHETALHCAAAAPYKPKKOICELLIRKAGINAKETKPLTPLHVAS 482

DB 171 TLDRGLDENHRDDAGWTPHMAA---FEGRHRLICEALIEQAGARTNEIDNDRIFILAS 227

QY 483 EKAHNDVVEVVVVKHAKVNALDNLQTSLHRAAYCGHILQTCRLLLSYGCDPNILSQGFT 542

DB 228 QEGHYDCVOILLENKSNIDQRYDGNALRVAALGHRDIVELFLSHGAD----- 277

QY 543 ALQMGNNVQQLLQGGISLGNSEADRLLEAKAGADVETVKLCIVQSNCRDIEGROST 602

DB 278 -----VNCKDADGRPT- 288

QY 603 PLHFAAGYNRVSVVEYLQHGADYHAKDKGGLVPLHNACSYGHYEVAELLVYKHGAVNVA 662

DB 289 -LYILALENQLTMABYFLLENGANVEASDAEGRTALHVSQWGHMEMVQVLYAHADVNA 347

QY 663 DLWKFTPLHEAAAKGVEICKLLLOLHGDADPTKKNRDGNTPLDLVKDGDTDIDLLRGAA 722

DB 348 DNEKRSALQSAAGQGHVKKVOLLIEHG-----A 375

QY 723 LLDAKAGKCLARVKKLLSPDNVNCRTQGRHSTPLHLAAGYNNLEVAEYLIQHGADVNAQ 782

DB 376 VVD-----HTCNOGATACLTAAQEGHIDVVQVLLLEHGDADPNHA 413

QY 783 DKGGLPLHNAASYGHVDVAALLIKYACVNAVDKWAFTPLHEAAKQGRQTQLCALL----- 838

DB 414 DQFGRTAMRVAANKHGSQIIKLEKYGA--SSLNGCSPSPVHTMEQKPLQSLSSKVSQSLT 471

QY 839 -----LAHG-----ADPTLKNQGGTPLDLVSAADDYSAL 867

DB 472 IKSNSSCGTGGDMQPSLRGLPNGPTHAFSPSPESPDSTVDROKSSUNSLSKSSKSSSL 531

QY 868 LTAAMPSPA-----LPSCYKQVNLGVRSFGATADALSSGSPSP 906

DB 532 RTTSSATATAQTVPIDSPHNLSTETEIQHSLPRSRSQ---SIVSPSTTQSLGQSHNSP 588

QY 907 SS-----LSAASSLDNLGSGFSELSSVSSVSGTEG--ASSLEKKEVPGVDFTSITQFVR 957

DB 589 SSEFEWSQVKPSLUKSKASKGKSENSAKSGSAGKAKQSNSSQPKVLEYEMTQFDR 645

RESULT 9

LATA_LATMA STANDARD; PRT; 1401 AA.

ID LATA_LATMA

AC P23631; Q25328;

DT 01-NOV-1991 (Rel. 20, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Alpha-latrotoxin precursor.

OS Latrotoxin mactans (Black widow spider).

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;

OC Araneomorphae; Entelegynae; Araneidae; Theridiidae; Latrodectus.

OX NCBI_TaxID=6924;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Tredecimquattus; TISSUE-Venom gland;

RX MEDLINE=91031994; PubMed=1977615;

RA Miyatkin N.I., Dulubova I.E., Chekhovskaya I.A., Grishin E.V.;

RT "Cloning and structure of cDNA encoding alpha-latrotoxin from black

widow spider venom.";

RL FEBS Lett. 270:127-131(1990).

RN [2]

RP PARTIAL SEQUENCE.

RC TISSUE-Venom;

RX MEDLINE=91362695; PubMed=1888339;

RA Volkova T.M., Galkina T.G., Kudelin A.B., Nazimov I.V., Grishin E.V.;

RT "Structure of tryptic fragments of a neurotoxin from black widow

spider venom.";

RL Bioorg. Khim. 17:437-441(1991).

CC -!- FUNCTION: Binds to high-affinity receptors that are localized in

the presynaptic plasma membrane of the nerve terminal. Binding of

toxin results in calcium influx into nerve terminals, activation

of synaptic vesicle exocytosis, and ATP depletion

-!- PTM: ALPHA-LATROTOXIN MAY BE PROCESSED AT ITS C-TERMINUS.

-!- SIMILARITY: CONTAINS 21 ANK REPEATS.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; X55009; CAA38753.1; -

DR PIR; S11527; S11527.

DR HSP; Q00420; LAMC.

DR InterPro; IPR002110; ANK.

DR Pfam; PF00023; ank; 20.


```

Db 400 -TPLTACKYNNTEAVKYLIESGCDTNDVMSFTPLNAVYVEREDTVKILLESQVDP 458
Qy 660 NVADLWKFTPLHEAARAGKYEICKLILLOHCAADPTKNNR-DGNTPLDLVKDQDDTDIQLLR 718
Db 459 NCDYKEVSPITQAIKRNKNKIKMLNAGIDIKPINECYGLHMLAALHN-----KDLL- 512
Qy 719 GDAALLDAAKKGLARVKKLSPDNNCRDQGRHSTPL-HLAAGYNLNLEVAEYLLQHGA 777
Db 513 -----KWLCTISEL-----EVNGVDD---HYVPLASYVAEISDIRIMELLIEKGL 555
Qy 778 DYN---AODKGGILPLHNAASHGVHVAALLI-----KYNACVNATDKWAF 821
Db 556 DLNKVTGPDETFTMTFSATSDLRKSIIDLLISOIAADEEFSECFKINKNNKIOTDKYLLR 615
Qy 822 PLHEAAQGRGTQCALLLAHGADPTLKQEGQTPDLVLS---ADDYSALLTAAMPSPAL- 877
Db 616 VYHEC-----KNQVSKM-----GETKLGDTMIDIYKNNRSIHVNFLARYAMQLSTID 664
Qy 878 -----PSCYKPKOVNGVR 890
Db 665 LREVPYRKYLELILNPAIKRHKILNAAK 693

RESULT 12
ANR3 HUMAN
ID ANR3 HUMAN STANDARD; PRT; 832 AA.
AC P50708; O96KH0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase ANKRD3 (EC 2.7.1.-) (Ankyrin repeat
domain protein 3) (PKC-delta-interacting protein kinase).
GN ANKRD3 OR DIK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Fetal kidney, and Fetal lung;
RA Shimizu N., Kudo J., Shibuya K.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudo J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordstark G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand B.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehtach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21."
RL Nature 405:311-319(2000).
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 10 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; AB047783; BAB56136.1; -.
DR EMBL; AP001743; BAA95526.1; -.
DR HSSP; P25963; LIKN.
DR Genew; HGNC:496; ANKRD3.
DR MIM; 605706; -.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00023; ank; 10.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00248; ANK; 10.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50088; ANK_REPEAT; 9.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW ANK repeat; Alternative splicing
FT DOMAIN 22..286 PROTEIN KINASE.
FT REPEAT 485..514 ANK 1.
FT REPEAT 518..547 ANK 2.
FT REPEAT 551..580 ANK 3.
FT REPEAT 584..613 ANK 4.
FT REPEAT 617..647 ANK 5.
FT REPEAT 651..680 ANK 6.
FT REPEAT 684..713 ANK 7.
FT REPEAT 717..746 ANK 8.
FT REPEAT 750..780 ANK 9.
FT REPEAT 782..811 ANK 10.
FT NP_BIND 28..36 ATP (BY SIMILARITY).
FT BINDING 51..51 ATP (BY SIMILARITY).
FT ACT_SITE 143..143 BY SIMILARITY.
FT VARSP_LIC 278..325 MISSING (IN ISOFORM 2).
FT CONFLICT 714..714 M -> V (IN REF. 1).
SQ SEQUENCE 832 AA; 91610 MW; 5D8FFFD5F04F7ECB CRC64;

Query Match 6.0%; Score 390.5; DB 1; Length 832;
Best Local Similarity 28.1%; Pred. No. 6.2e-16;
Matches 141; Conservative 58; Mismatches 176; Indels 127; Gaps 13;

Qy 368 SLLLSYGADPTLLNCHNKSADLDLPTPLKRLAYEFKGHSLLQAAAREADVTRIKKHLSL 427
Db 434 SLSLSFEREP-----STSLGTITDVQKK-----LYDAIVSGDTSKLMKIL-- 474
Qy 428 EMVNEKHPQTHTALHCAAA----SPYPRKQICELLRLRGANINEKTEFLPLHVASE 483
Db 475 -----QPQDVLALDSGASLLHLAVEAGOECAKWLNNANPNLSNRGRGSTPLHMAVE 528
Qy 484 KAHNDVVEVVKHEAKVNALDNLGQTSLHRAAYCGHLOTCTRLLSYGCDDNIIISLOGFT- 542
Db 529 RVRGVWELLARKTSVNAKDEQDQWTLHPAQNGDESSTRLLEKNASVNEVDFEGRT 588
Qy 543 ---ALQMGNEH-VQOLLQEGISLGNSEADQRLLEAAKAGDVETVKLCTVQSVNCRDIEG 598
Db 589 MHVACQHGGENIVRIILRRGVV-----SIQ 615
Qy 599 ROS-TPLHFAAGYNRVSVVEYLL-QHGADYHAKDKGLVPLHNACSYGHYEVAELLVKG 656
Db 616 KDAPLPHVAAWQGHLPVKKLAKQPCVSNQATLDGRTPLHLAAQGRHYRVARILDL 675
Qy 657 AVNVADLWKFPLHEAARAGKYEICKLILLOHCAADPTKNNR-DGNTPLDLVKDQDDTDIQL 716
Db 676 SDNVNCSLLAQPTPLHVAATGHTSTARLLHARGAGKAMTSDGYTAUHL----- 724
Qy 717 LRGAALLDAAKKGLARVKKLSPDNNCRDQGRHSTPLHLAAGYNLNLEVAEYLLQHG 776
Db 725 -----AARNGLHATVK-----LLVEEK 741

```



```
DR Pfam; PF00023; ank: 10.
DR SMART; SM00248; ANK: 6.
DR PROSITE; PS50088; ANK_REPEAT; 7.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW Hypothetical protein; Repeat; ANK repeat.
FT REPEAT 40 69
FT REPEAT 144 173
FT REPEAT 177 206
FT REPEAT 210 239
FT REPEAT 272 302
FT REPEAT 306 336
FT REPEAT 340 370
FT REPEAT 374 403
FT REPEAT 407 437
FT REPEAT 441 471
FT REPEAT 473 502
FT REPEAT 571 602
SQ SEQUENCE 668 AA; 76200 MW; F3C5BA076023791B CRC64;

Query Match          5.9%; Score 380.5; DB 1; Length 668;
Best Local Similarity 27.3%; Pred. No. 1.8e-15;
Matches 153; Conservative 95; Mismatches 192; Indels 121; Gaps 26;

QY 326 CSYGHVEVTELV-----KHGA--CVNMDLWQPTPLHEAASKNRVEVCSLLSYGADPT 378
   | : | : | | | | | : | : | : | : | : | : | : | : | : | : | : |
Db 11 CRGNYEILSAITNYNLHKHGANOCEN--ESIPFTAIHQALQLRQIDIVKELIQ--QNP 66

QY 379 LL-----NCHNKAIDAPTOLKERLAYEPKGLSLQAAREADYTRIKKHLSELMVNF 434
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 67 LIVYDHRHRSITLHTICITPNVMD-----IVISLTVDCDILDIKYAST-ILN-KH 115

QY 435 POTHETALHCAAAASPYPKRRQICELLRLKRGANINE---KTKEFLTPLHVAASEKAHND-- 488
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 116 -----KIGEAETHVLKEGISGNEISYNKINKSIEMKLIKERTQQDEL 158

QY 489 -VVEVVVVKHEAKVNALNLSGTSLHRAAYCYGHLQTCRLLLSYGCDPNILISOGFTALQ-- 545
   | : | : | | | | | : | : | : | : | : | : | : | : | : | : | : |
Db 159 LIAEMLLKKGIDVNAKDVCYCTPIHYAERGNTRKMNLLLSYGADVNIITLDDLSVLEYA 218

QY 546 MGENVVOQLLEGISIGN--SEADROLLEAAKAGDVEVTKKLC--TVQSVNCRDIEGRQSTP 603
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 219 VDSKNIDITKAITDNRSNINKNDLSLKAITNTOLETSLLLYDSGFSVNSIDV--YKNT 276

QY 604 LHFAAGYNRVS-VVEYLLQHGADVHAKDGLVPLHNACSYGH-YEVAELVKKHGVNV 661
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 277 LHVAQAPSLRSLVPKLLERIGIDVNAKNIKGETPLYLMKNGYDTNIRTLIMRGADVNA 336

QY 662 ADLWKFTPLHEAAAKGYEICKL-LLOHGADPTKKNRDNPTPLDLVKDGDITDIQDLLRGD 720
   | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 337 ADSLYITPLHQASTLDYIKDTVITLLELGA----- 366

QY 721 AALLDAAKKGCLARVKKLSPPDYNCRDTQGRHSTPLHLAAGYNLNLEVAEYLLQHGADV 780
   | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 367 -----NVNARDYCDK--TPIHYAARNVNVLIINTLLDYGADIE 402

QY 781 A-QDKGGLIPLHNAASYG---HVDVAALLIKYNACVNATDKWAFITPLHEAAQKG-RTQLC 835
   | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 403 ALSQKIGTV-LH-PALYGTNPYMSVKT-LIDRGANVNSKNKYLSTPLHYACKKNCKPEVI 459

QY 836 ALLLAHGADPTLKNQEGOTPL 856
   | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 460 KMLLDNGADVNAIRNQYPL 480

RESULT 15
ID ASB2_HUMAN STANDARD; PRT; 587 AA.
AC Q96Q27; Q9NSU5; Q9Y567;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin repeat and SOCS box containing protein 2 (ASB-2).
GN ASB2.
```

```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND INDUCTION.
RX MEDLINE=21450655; PubMed=11566180;
RA Kohroki J., Fujita S., Itoh N., Yamada Y., Imai H., Yumoto N.,
RA Nakanishi T., Tanaka K.;
RT "ATRA-regulated Asb-2 gene induced in differentiation of HL-60
RL leukemia cells.";
RL FEBS Lett. 505:223-228(2001).
RN [2]
RP SEQUENCE FROM N.A., AND INDUCTION.
RX MEDLINE=21634875; PubMed=11682484;
RA Guibal F.C., Moog-Lutz C., Smolewski P., Di Giola Y.,
RA Darzynkiewicz Z., Lutz P.G., Cayre Y.E.;
RT "ASB-2 inhibits growth and promotes commitment in myeloid leukemia
RL cells.";
RL J. Biol. Chem. 277:218-224(2002).
RN [3]
RP SEQUENCE OF 31-587 FROM N.A.
RC TISSUE=Uterus;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 42-587 FROM N.A.
RX MEDLINE=20564172; PubMed=11111040;
RA Kile B.T., Viney E.M., Willison T.A., Brodnicki T.C., Cancilla M.R.,
RA Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA Alexander W.S.;
RT "Cloning and characterization of the genes encoding the ankyrin repeat
RL and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";
RL Gene 258:31-41(2000).
CC !- FUNCTION: Not known. Its expression induces growth inhibition and
CC chromatin condensation.
CC !- INDUCTION: By all-trans retinoic acid (ATRA).
CC !- SIMILARITY: CONTAINS 12 ANK REPEATS.
CC !- SIMILARITY: CONTAINS 1 SOCS BOX DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB056723; BAB64532.1; -
CC EMBL; AJ251238; CAC17765.1; -
CC EMBL; AL137735; CAB70899.1; -
CC EMBL; AF159164; AAD45345.1; -
CC Genew; HGNC:16012; ASB2.
CC MIM; 605759; -
CC InterPro; IPR002110; ANK.
CC InterPro; IPR001496; SOCS.
CC Pfam; PF00023; ank; 11.
CC PROSITE; PS50088; ANK_REPEAT; 9.
CC PROSITE; PS50297; ANK_REPEAT; 1.
CC PROSITE; PS50225; SOCS; 1.
KW ANK repeat; Repeat.
FT REPEAT 56 85
FT REPEAT 89 119
FT REPEAT 123 152
FT REPEAT 156 185
FT REPEAT 189 218
FT REPEAT 222 251
FT REPEAT 255 284
FT REPEAT 288 317
FT REPEAT 320 349
FT REPEAT 362 391
FT REPEAT 392 421
```

```
FT REPEAT 428 456 ANK 12.
FT DOMAIN 533 587 SOCS BOX.
FT CONFLICT 224 224 MISSING (IN REF. 4).
FT CONFLICT 460 460 MISSING (IN REF. 4).
SQ SEQUENCE 587 AA: 65083 MW: 866A2B6010C41F9 CRC64;

Query Match
Best Local Similarity 23.7%; Score 357; DB 1; Length 587;
Matches 146; Conservative 80; Mismatches 215; Indels 174; Gaps 18;

QY 102 LPEACRNGDVERVKRLVTEKVNRSRTAGRKSTPLHFAAGFGKRDVVEY--LQNCANVQ 159
Db 12 LFHS-----SAPSRSTAPPESPAPMGLFQGVMMQYSSSLFKTSOLA 56
QY 160 ARDDGGLIPLHNACSFGEVNVNLLRHG---ADPNARDNNVYTPLEHAAIKGKIDVCIV 216
Db 57 PAD-----PLIKAKDGDDEALKTMIKEGKNLAEPN-KEGW--LPLHEAAYGQVG-CLK 107
QY 217 LLOHGAEPTRNTDRTALDADPSAKAVLTGEYKKDELLESARSNGEEKMMALLTPLNV 276
Db 108 VLQRAYPGTI---DORT---LQETAVYLATCRGHLDCLLSLLOACAEPDI----- 152
QY 277 NCHASGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLVPLHNACSYGHYEVTTEL 336
Db 153 ----SNKSRETPLYKACERKNAEAVKILVQHNADTNHRCNRGWTALHESVSRNDLEVMQI 208
QY 337 LVKHGACVAMDLWQFTPLHEAASKNRVEVCSSLLSYGADPTLLNCHNKSAIDLAFTPOL 396
Db 209 LVSGGAKVESKNAYGITPLFVAASQGLEALREFLAKYGADINTQASDNASA----- 259
QY 397 KERLAYEFKGHSLQAAAREADVTRIKKHLSELMVNFKHPQTHETALHCAAASPYPKRKQI 456
Db 260 -----LYEACKN-----EHEEV 271
QY 457 CELLLRKGANINEKTEFLTPLVHASEKAHNDVVEVVVHEAKVNALDNLGOTSLHRAAY 516
Db 272 VEFLLSQGADANKTKDGLPLHLIASKKNYRIVOMLLPVTSTR-IRRSVGSPLHAAE 330
QY 517 CGHLOTCLLLSYGCDPNIISLQGTALQMGNEVQQLQEGISLGNSEADROLLEAACA 576
Db 331 RNHDEVLEALLSAREFVN----- 348
QY 577 GDVETVKLCITVOSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVP 636
Db 349 -----TPLAPERARLYEDRRSSALYFAVNNNNVYATELLLOHGAD---PNRDVISP 396
QY 637 LHNACSYGHYEABLLVKGAVNVADLWKFTPLHEA-----AKGYETCKLLQLHG 689
Db 397 LLVAIRHGCLRTMQLLLDHGANIDA-----YIATHPTAFPATIMFAMKCLSLKFLMDLIG 451
QY 690 AD--PTKKNRDGNTP 702
Db 452 CDGEPFCFCLYNGP 466
```

Search completed: February 12, 2003, 03:33:06
Job time : 20.3675 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 03:26:09 ; Search time 46.6325 seconds
(without alignments)
5478.980 Million cell updates/sec

Title: US-09-843-159B-4
Perfect score: 6464
Sequence: 1 RCSARRGAAGGQAGRGARV.....AYPEYLITYQIMRPEGMVDG 1240

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4005	62.0	1181	5 Q9VBP3	Q9vbp3 drosophila
2	3997	61.8	1181	5 Q9XZ37	Q9xz37 drosophila
3	904	14.0	1136	6 Q9N180	Q9n180 bos taurus
4	893	13.8	1726	11 Q8VC68	Q8vc68 mus musculus
5	892.5	13.8	4377	4 Q12955	Q12955 homo sapien
6	891	13.8	1943	11 Q61307	Q61307 mus musculus
7	888.5	13.7	1762	11 Q88521	Q88521 rattus norv
8	885.5	13.7	2622	11 Q70511	Q70511 rattus norv
9	880	13.6	1856	4 Q99407	Q99407 homo sapien
10	878.5	13.6	1719	4 Q13768	Q13768 homo sapien
11	876	13.6	1848	11 Q61302	Q61302 mus musculus
12	868.5	13.4	1549	5 Q9V4B1	Q9v4b1 drosophila
13	866.5	13.4	1549	5 Q24241	Q24241 drosophila
14	858	13.3	843	11 P97582	P97582 rattus norv
15	858	13.3	1809	5 Q17487	Q17487 caenorhabdi
16	858	13.3	1867	5 Q17486	Q17486 caenorhabdi

17	858	13.3	2039	5	Q17489	Q17489 caenorhabdi
18	854	13.2	1815	5	Q17488	Q17488 caenorhabdi
19	854	13.2	6994	5	Q17343	Q17343 caenorhabdi
20	851	13.2	1786	5	Q17344	Q17344 caenorhabdi
21	848.5	13.1	1159	5	Q9NCP8	Q9ncp8 drosophila
22	827.5	12.8	2443	5	Q9VSA2	Q9vsa2 drosophila
23	817	12.6	1009	5	O8SWY2	O8swy2 drosophila
24	657	10.2	791	4	Q9UFA4	Q9ufa4 homo sapien
25	635.5	9.8	1619	5	Q9NGR8	Q9ngr8 drosophila
26	629.5	9.7	1088	4	Q13484	Q13484 homo sapien
27	627	9.7	1913	5	Q9GRV5	Q9grv5 caenorhabdi
28	609	9.4	1453	5	Q9VMR4	Q9vmr4 drosophila
29	590	9.1	1599	11	Q9NHH0	Q9nhh0 mus musculus
30	563.5	8.7	934	16	Q83807	Q83807 treponema p
31	557	8.6	1486	4	O8TEF1	O8tef1 homo sapien
32	554.5	8.6	1098	11	Q61304	Q61304 mus musculus
33	552	8.5	1411	5	O02989	O02989 latrodectus
34	539.5	8.3	1395	5	Q9XZC0	Q9xzco latrodectus
35	539	8.3	833	4	Q96186	Q96186 homo sapien
36	530.5	8.2	1282	5	O8T4F8	O8t4f8 drosophila
37	530.5	8.2	2119	5	Q9VAU5	Q9vau5 drosophila
38	530.5	8.2	2119	5	Q8T9G9	Q8t9g9 drosophila
39	527	8.2	743	4	Q8TB46	Q8tb46 homo sapien
40	525.5	8.1	4001	5	Q8WRQ7	Q8wrq7 drosophila
41	524	8.1	1188	4	Q9H288	Q9h288 homo sapien
42	509	7.9	1021	13	O8UVC1	O8uvcl brachydantio
43	507.5	7.9	1058	16	Q9PCW4	Q9pcw4 xylella fas
44	505.5	7.8	686	4	Q9NXF0	Q9nxf0 homo sapien
45	495	7.7	861	5	Q8WP35	Q8wp35 geodia cydo

ALIGNMENTS

RESULT 1

Q9VBP3
ID Q9VBP3 PRELIMINARY; PRT: 1181 AA.
AC Q9VBP3;
DC Q9VBP3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE BCDA:LD22548 protein.
GN BCDA:LD22548 OR CG4719 OR CG17487.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwag C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., M.G.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhong X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Venter G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*";
RL Science 287:2185-2195(2000).
DR EMBL: AE003752; AAF56487.1; -.
DR HSP: Q00420; IAWC.
DR FlyBase: FBgn0027508; BcDNA:LD22548.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR001660; SAM.
DR Pfam: PF00023; ank; 17.
DR Pfam: PF00536; SAM; 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 15.
DR SMART: SM00454; SAM; 1.
DR PROSITE: PS50088; ANK_REPEAT; 14.
DR PROSITE: PS50297; ANK_REPEAT; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1181 AA; 127944 MW; EB6F0F94E6A3D504 CRC64;

Query Match 62.0%; Score 4005; DB 5; Length 1181;
Best Local Similarity 65.3%; Pred. No. 7.4e-242;
Matches 772; Conservative 166; Mismatches 189; Indels 56; Gaps 13;

QY 91 ABAVPEPAARELFACRNGDVERVKRLVTEPKVNSRDTAGKSTPLHFAAGFGKDVVEY 150
DB 17 AVMPDP-LRELFPEACKTGEIAKVLITPOTNARDTAGKSTPLHFAAGFGREVEF 75
QY 151 LQONGANVOARDGGIPLHNACSFHAEVNVLLRRGADPNARDNNYTPLEAAIKGK 210
DB 76 LNSGASIQACDEGGLPHLNCSCFHAEEVRLKLGACSPNTDNNYTPLEAAASKG 135
QY 211 IDVCIVLQHGAEPTINTDGRALTADPSAKAVLFGYKDELLSAGNNEKMMAL 270
DB 136 VDVCALLQHGANTIRNSQKTPLEADEATREVLTEGYRKDELLEAARSAGEDRLIAL 195
QY 271 LTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDGLVPLHNACSYGH 330
DB 196 LTPLNVNCHASDGRKSTPLHLAAGYNRIGIVETILLANGADVHAKDGLVPLHNACSYGH 255
QY 331 YEVTLLVKGACVAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSATDL 390
DB 256 FDTVKLLIQGANVANDLWAFPTPLHEAASKSRVEVCSLLLSRGADPTLLNCHSKSAIDA 315
QY 391 APTPOLKERIAYEFKSHLLQAAAREADVTRIKKHLISLEWNVNFKHPQTHETALHCAASPY 450
DB 316 APTRELRIATAFEYKGHCLLDACKKCDVSRAKLVCAEIVNFVHPYTGDTPLHLAVVSPD 375
QY 451 PKRQICELLRKGANINTEKTEFLTPLVHASEKAHNDVVEVVKHAKVNALNGLGTS 510
DB 376 GKRQKQMLLETRKGLLNENKNAFLTPHLAAEALLHYDAMEVLKQKAGNALDSLOQTP 435
QY 511 LHRAAYCGHLOTRCLLSYCDPNIIISLOGFTALQMGNNVQQLLOEGISLGNSEADROL 570
DB 436 LHACA--RDQAVRLLSYADNTNIVSLEGLTAQIALSDSVLKLLK-----NPPDSETHL 488
QY 571 LEAAKAGDVETVKKLCITVQ--SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHCADVHA 628
DB 489 LEAAKAGDLDVTRVRLNNPISVNCRLDGRHSTPLHFAAGFNRPVVPVQFLEHGAEYVA 548
QY 629 KDKGGLVPLHNACSYGHYEVAELLVKHGVNVNADLWKFPTPLHEAAAKGKYEICKLILQH 688

DB 549 ADKGLVPLHNACSYGHYEVELLYVKHGANVNSDLWKFPTLHEAAAKGYDICKLILK 608
QY 689 GADPTTKNRDGTPLDLYKVGDDTDIDQLLRGDAALLDAAKGCGLARVKKLSSPDNVCRD 748
DB 609 GADPMKNRDGTADPADLYKESDHDVAELLRGPSSALLDAAKGNLARVORLVPESINCRD 668
QY 749 TQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDAALLIKY 808
DB 669 AQGRNSTPLHLAAGYNNFECAEYLLLENGADVNAQDKGLIPLHNASSYGHDLTAALLIKH 728
QY 809 NACVNATDKWAFPTPLHEAAAKQRTQLCALLAHGADPTLKNQGPOTPLDVSADDSALL 868
DB 729 KTVVNATDKWFTPLHEAAAKQRTQLCULLAHGADAYMKNQGPOTPIELATADDVKCLL 788
QY 869 TAAMPSS---ALPSCYKPVQLNGVRSPGATADALSSGSPSSPSS--LSAASSLDNLSSGF 922
DB 789 QDAMATSLSQOAL-SASTQSLTSSSPADPATAAAGPTSSSSSAILSPPTETVLLPTGA 847
QY 923 SELSSV-----VSSS-----GTEGASSLEKKEVPGVD--FSITQFVRNLGLEH 963
DB 848 SMILSVVPVPLPLSSSTRISPAQCAEANGAEGSSDDL--LPDADTITNVSGFLSSQOLHH 905
QY 964 LMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLKGVRLISSQOGLNPLYTLNTS 1023
DB 906 LIEUFEREQITLDILAEMGHDDLKQGVSAFGYFRHKLKGLIAQLRS-----TT 953
QY 1024 GSG-----TILIDLSPDKFQSVVEEMQSTVREHRDGGHAGGIFNRYNLIKIQKVCNK 1077
DB 954 GIGNNVNLTLLVOLLPPDKFEVAVEEMQATIREHRDNGOAGGYTRYNIIRVKVONR 1013
QY 1078 KLWERYTHRRKEVSEENHNANERMLFHGSPFVNAILHKGDFDERHAYIGMFGAGYFAE 1137
DB 1014 KLWERYTHRRQRIAEENFLQSNERMLFHGSPFVNAILVQRFDERHAYIGMFGAGYFAE 1073
QY 1138 NSSKNQVYVGGTGGCGPVHDKRSCYICHRQLFCRVTLLKGFLOFSAMKMAHSPGHH 1197
DB 1074 HSSKNQVYVGGTGGCGPVHDKRSCYICHRQLFCRVTLLKGFLOFSAMKMAHSPGHH 1133
QY 1198 SVTGRSVNGLAALAEVYIRGEQAYPEYLITYOIMRPEGMVDG 1240
DB 1134 SVVGRPSAGGLHFAEVYVYRGQSYPEYLITYOIVKPDSSSG 1176

RESULT 2
Q9XZ37
ID Q9XZ37 PRELIMINARY; PRT: 1181 AA.
AC Q9XZ37;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical 127.9 kDa protein.
GN BCDNA:LD22548 OR CG4719 OR CG17487.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID:7227;
RP SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenhoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svirkas R.R., Weinburg T.,
RA Celniker S.E.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF132196; AAD34784.1; -.
DR HSP: Q00420; IAWC.
DR FlyBase: FBgn0027508; BcDNA:LD22548.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR001660; SAM.


```
Db 211 IAAHYENLVAQLLILRGASVNFPTQNGIITPLHIASRRGNVIMVRLLLDRGAQIETRTKD 270
QY 318 DLVPLHNACSYGHEVTELLVKGACVAMDLWQFTPLHEAASKNRVEVCSSLISYGA-- 375
Db 271 ELTPLHCAARNGHLRISEIILLDHGAPTOAKTKNGLSPIHMAAQGDHLDVCLLLOYNAEI 330
QY 376 -DPTL-----LNC-HNKA---IDLAPTQPKERLAYEFKGHSLLQAREADVTRI 421
Db 331 DDITDLHLTPLHVAACHGHHRAKAVLLDKGAKPNR-----ALNGFTPLHTACKKNHIV 385
QY 422 -----KKHLSLEWVNFKHPOTHETALHCAASPYPKKQICELLIRKGANINEKTELT 476
Db 386 MELLKMGASIDAVT-----ESGLTPLHVSFMPGP---PIVKSLLQREASPNVSKVET 438
QY 477 PLHVAEKANDVVEVVVKEAKVNALDNLGOTSLHRAAYCGHLQTCRLLISYGCDPNII 536
Db 439 PLHMAARAGHTEVAKYLDQWAKVNAKAKDDOTPLHCAARIGHTNMVKLLLENANPNLA 498
QY 537 SLOGFTALOMGNVVOQLLOEGISLGNSEADRLQLEAAKAGDVETVKKLCTVQ-SVNCRD 595
Db 499 TTAGHTPLHI-----AAREGHVETALALLEKEASQTCMT 532
QY 596 IEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHEVAELLVKH 655
Db 533 KKG--FTPLHVAAYKYGKVRMAELLEDHDAHPNAAGKSGLTPLHVAVHHNHLDVVRLLPR 590
QY 656 GAVVNVADLWKFTPLHAAAAGKYEICKLLLOHGADPTKKNRDNLTPLDL-VKDGDTDIQ 714
Db 591 GGSFHPALNGYTPPLHTAAKONQLEVARSLLOYGGSNAESVQGYTPHLHAAQEGHAEV 650
QY 715 DLRGDAALLDAKKGCLARVKKLSSPDNNVNCRTQGRHSTPLHLAGYNLEVAEYLLQ 774
Db 651 ALL-----LSKANGNLGNKSG--LTPHLVAGQEGHIPVADVLIK 688
QY 775 HGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWFTPLHAAQKGRQL 834
Db 689 HGVTVDATTRMGYTPPLHVAHYGNIKLVFLLQHKADVNAKTKILGYSPLHQAQOQHGTDI 748
QY 835 CALLAHGADPTLKNBQGTPLDL-----VSADDVSALITAAAMPSPALPCYK---PQV 885
Db 749 VTLHLKHGASPNVSSNGTTPATAKRLGYISVTDVLKVVTDETTVQLITDKHRMSPPET 808
QY 886 LN-----GVRSPGATA-----DALSSGSPSP-----SS 908
Db 809 VDEILDVSEDEGLGPRABRPDARDVDEKELLDVFPKLDQVVESPAIPRIPCVTPTV 868
QY 909 LSAASSLDNLGSEFSELSVSSSGTEGASLEKKEVPGVDFTSQTFVRNLGLEHLMDF 968
Db 869 VIRAEPDQVSKYEDESLIPSSPATETSDNISPVASGVHTGFLVSF----- 916
QY 969 FREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVE-----RLISGQQGLN 1015
Db 917 -----VDAGGSMRSGRHNGLRVVPVPPRCAATPRTICRLVRPKLPT 959
QY 1016 PYLTLTSGGTLLIDLP-----DDKEFQSVSEEMOSTVVRHR 1054
Db 960 PPLAAEEGLASRVIALGPTGVQVFLSPVIVEIPHAFASQGRGDRVLRSENGSVWKEHK 1019
QY 1055 DGHGAGGIFNRYNTLKTOKVCN 1076
Db 1020 -----NRYGESYLDQILN 1032

RESULT 4
ID Q8VC68 PRELIMINARY; PRT; 1726 AA.
AC Q8VC68;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 188.2 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary Gland;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC021657; AAH21657.1; -.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000488; Death.
DR InterPro: IPR000906; ZUS.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; ZUS; 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 22.
DR SMART: SM00005; DEATH; 1.
DR SMART: SM00218; ZUS; 1.
DR PROSITE: PS50088; ANK_REPEAT; 21.
DR PROSITE: PS50297; ANK_REPEAT; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KW Hypothetical protein.
SQ SEQUENCE 1726 AA; 188239 MW; 430148B67E00B1CA CRC64;

Query Match 13.8%; Score 893; DB 11; Length 1726;
Best Local Similarity 28.3%; Pred. No. 8.1e-47;
Matches 337; Conservative 167; Mismatches 497; Indels 188; Gaps 38;

QY 68 AVAAARIMSGRCAGGGACACASAAAEVPAARELEFEACRNGDVERVKRLVTEPKVNSRD 127
Db 10 AKPAHRKRKKGSNDANASYLRAARAGHLEKA----LDYIKNG-----VDVNICN 54
QY 128 TAGRSTPLHFAAGFGKDKVVEYLLQNGANVOARDGGGLIPLHNACSFHAEVYNLLRH 187
Db 55 QNGLNA--LHLASKEGHEVVEVSELLQREANVDAATKKGNTALHIASLAGQEVVKLVTN 112
QY 188 GADPNARDNNWYTPHLEAAIKGIDYICIVLLOHGAETPTIRNTDGTALDLA-----DPSA 242
Db 113 GANVAQSNGTGYPMYAAQENHLEVRFLDNGASQSLATEDGFTPLAVALQOQHQQV 172
QY 243 KAVLTGEYK--KDELESARSGNEEKMAALLTPLNVNCHASDGRKSTPLHLAGYNRVK 299
Db 173 SULLENDTKGKRLPALHIAARKDDTKAAALLQNDTNADVESKSGFTPLHIAAHYGNIN 232
QY 300 IVQLLQHGADVHAKDKGDLVPLHNACSYGHEVTELLVKGACVNAAMDLMWQFTPLHEAA 359
Db 233 VATLLNRAAAVDFTARNDTPLHVAASKRGANVMKLLLDGAKIDAKTRDGLTPLHCGA 292
QY 360 SKNRVEVCSSLISYGADPTLLNCHNKSAIDLAFTPOLKERLAYEFKGHSLLQAREADV 419
Db 293 RSGHEQVVEMLDRSA-PILSKTN-----GLSPLHMATQGD-- 328
QY 420 RIKKHLSLEWVNFKH-----PQTHETALHCAASPYPKKQICELLIRKGANINEKTK 472
Db 329 ----HLNCVOLLQHNVPVDDVTNDYLTALHVAACHGCHYK---VAKVLLDKKASPNKAL 381
QY 473 EFLTPLHVAASEKAHNDVVEVVVKEAKVNALDNLGOTSLHRAAYCGHLQTCRLLISYGD 532
Db 382 NGFTPLHIAACKNRIRVMELLKHGASIQAVTESGLTPIHVAAFMGHVNIVSOLMHGAS 441
QY 533 PNIISLOGFTALOMG-----NENVQOLLQEGISLGNSEADRO--LLEAAKAGDVETVKKL 585
Db 442 PNTTNVGRGTALHMAARSQAEVRYLVQDGAQVAKAKDDQTPPLHLSARLGRADIVQOL 501
QY 586 CTV-OSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYG 644
Db 502 LOQGASPNAAITSG--YTPHLHAAAREGHEDVAAFLDHGASLSITTKGFTPLHVAAYK 559
QY 645 HYEVAELIVKHGAVNVNADLWKFTPLHLEAAKGYEICKLLLOHGADPTKKNRDNLTPLD 704
Db 560 KLEVASLLQKSPADPAACKSGLTPLHVAHYDNQKVALLLLDQGASPHAAAKNGYTPLH 619
QY 705 L-VKDGDTDIQDLL-----RGDAALLDAAKKGLARVKKLSSPD-VVNCRDQ 750
```


Db 620 IAANKNOMDIATSLLEYGADANAVTRQGIASVHLLAQEGHVDVMSLLSRNANVNLNKS 679
QY 751 GRHSTPLHLAAGYNNLEVAEYLLQHGADYNAQDKGLIPLHNAASGYCHVDVAALLIKYNA 810
Db 680 G--LTPHLAAGQEDRVNAEVLVNOGAHVDAOTKMGTTPPHVCHGYGNKIVNFFLQHS 737
QY 811 CVNATDKWFTPLHEAAQKRTOLCALLAHGADPTLKNQEGTPTLDLVS-----ADD 863
Db 738 KVNAKTKNGYTALHQAQOQTHIINVLQNNASPNELTVNGNTALAIARRLGIYSVVD 797
QY 864 VSALLTAAMPSPNALPCYK---POVLNGV--RSPGATADALS-----GP----- 903
Db 798 LKVTTEIIMTTTITTEKHKMNPETMNEVLDMSDDEGDAITGDTDKYLGPDOLKELGDD 857
QY 904 SSPSSLSAASLDNLSGSELSVSVVS-----SSGTEGASSLEKKEVPVGVDFSIIOFVN 958
Db 858 SUPAEGYGVFSLGARSASLSFSSDSRYTLNRSSVARDMSMTEELLVPSKEQHLT-FTRE 916
QY 959 LGLEHLMDFEREQITLDVL-----VEMGHKELKEIGINAYG-----HRHKLII--- 1001
Db 917 FDSDSLRLH-YSWAADTLDNVLNVSSPVHSGF--LVSMFVDARGGSMRSGRHHGMRIIIP 973
QY 1002 -----KGVRLISQOGLNPYLTLNTSGSGTILIDLSPDDKEFQS-VEEEMQSTVREH 1053
Db 974 RKTAPTTRITCLVRKHLKIANPPMVVEGEGLASRLVEMGPAGAQLGPIVIVEIP----- 1027
QY 1054 RDGHAGGIFRNYLNLIKQKCNKLWE--RYTHRKVEYSE-----ENHNHANERMLPHG 1106
Db 1028 ---HFGSMGRKERELIVLRSENGETWKEHQFDSKNEDLAELLNGMDELDSPPEL---G 1080
QY 1107 SPFVNAIHKGFDERHAYIGMGFGAGIYFAENS---SKSNQYVYIGGG 1152
Db 1081 TKRICRIITKDPQ-----YFAYVSRIKQESNQ--IGPEGG 1114

RESULT 5
Q12955 PRELIMINARY; PRT; 4377 AA.
ID Q12955
AC Q12955;
DC 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Ankyrin G.
GN ANK-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN STEM;
RX MEDLINE=95136209; PubMed=7836469;
RT Kordeli E., Lambert S., Bennett V.;
RT "AnkyrinG. A new ankyrin gene with neural-specific isoforms localized
RT at the axonal initial segment and node of Ranvier.";
RN J. Biol. Chem. 270:2352-2359(1995).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN STEM;
RA Carpenter S.S.;
RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U13616; AAA64834.1; -;
DR HSSP; P55273; 1B18.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.

DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 4377 AA; 480395 MW; F42379E55768B684 CRC64;

Query Match 13.8%; Score 892.5; DB 4; Length 4377;
Best Local Similarity 30.7%; Pred. No. 3.7e-46;
Matches 277; Conservative 126; Mismatches 365; Indels 135; Gaps 21;

QY 136 LHFAAGRGKDDVVEYLLONGANYOARDGGLIPLHNACSFGEHAEVNNLLRHGHGADPNARD 195
Db 78 LHLASKEGHEVVEVSELLQREANVDAATKKNTALHTASLAGQAQAEVVKVLTNGANVNAQS 137
QY 196 NNNYTPPLHEAAIKGKIDVCLVILLOHGAETPIRNTDGTALDLA-----DPSAKAVLTGEY 250
Db 138 QNGFTPLYMAAQHNEHLEEVVKFLDNGASOSLATEDEGFTPLAVALQOQHDQVSLLENDT 197
QY 251 K---KDELLESARGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQ 307
Db 198 KGKVRPLPALHIAARKDDTKAAALLLQNDNNADVESKSGFTPLHIAAHYGNINATLLNR 257
QY 308 GADVHAADKDDLPLHNACSYGHEVTELLYKHGACYNAMDLMQFTPLHEAASKNRVEVC 367
Db 258 AAADVFTARNITPLHVASKRGNNANVKLLDRGAKIDAKTRDGLTPLHCGARGHEQVV 317
QY 368 SLLSYGADPTLLNCNKNSAIDLAPTQPKERLAYEFKSHLSLQAAREADVTRIKKHL 427
Db 318 EMLLDRAA-PILSKTKN-----GLSLPHMATQGD-----HLNC 349
QY 428 EMVNFKH-----PQTHETALHCAASPYPRKROICEILLLRKGANINEKTEFLTPLHV 480
Db 350 VQLLLQHNVPDVTNDYTLALHVAACHGK---VAKVLLDRKANPNKALNGFTPLHI 406
QY 481 ASEKANDVVEVVKHEAKVNALDNLGOTSLHRAAYCGHLQTCRLLSYGCDPNIIISLQ 540
Db 407 ACKKNRIKVNELLKHGASQAVTESGLTPIHVAEFMGHVNIVSOLMHHGASPNNTNVRG 466
QY 541 FTALQMG-----NENVOOLLOEGISLGNSEADRO---LLEAKAGADVETVKKLCV-OSVN 592
Db 467 ETALHMAASGQAEVVRVYLVQDGAQVEAKAKDDOTPLHISARLGKADI VOQLLQOQASPN 526
QY 593 CRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHEVAELL 652
Db 527 AATTSG--YTPLHSAREGHEDVAFLDHDGASLSITTKGFTPLHVAAYKGLVANLL 584
QY 653 VKHGVVNVADLWKFTPLHEAAAKGYETCKLLQLQHGADPTKKNRDGNTPLDL-VKDGDT 711
Db 585 LOKSASPDAAGKSGLTPLHVAAYHYNQKVALLLDQGASPHAAAKNGYTPHIAAKKNOM 644
QY 712 DTIDLL-----RGDAALLDAAKGCLARVKKLSSPD-VNCRDTQGRHSTPLH 758
Db 645 DIATLLEYGADANAVTROGIAVSHLAQEGHVDVMSLLGRNANVNLNKS- -LTP 702
QY 759 LAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASGYCHVDVAALLIKYNAVNATDKW 818
Db 703 LAAQEDRVNVAEVLVNOGAHVDAOTKMGTPLHVGHYGNKIVNFFLQHSKAVNAKTN 762
QY 819 AFTPLHEAAQKRTOLCALLAHGADPTLKNQEGTPTLDI-----VSADDVSALIT-AA 871
Db 763 GYTPLHQAQOQTHIINVLQNNASPNELTVNGNTALGTARRLGIYSVVDTLKIVTEET 822
QY 872 MPSPALPCYK---POVLNGVRSFGATADALS GSPSPSSLSAASSLDNLSGSELSV 928
Db 823 MTTTIVTEKHKMNPETMNEVL-----MSDDEVKANAPEMLSDG----- 863
QY 929 VSSSGTEGASSLEKKEVPVGVDFSIIOFVRNLGLEHLMDFEREQ-ITLDVLEMGHKE 987
Db 864 -----EYISDVEEGEDAMTGTDKYLGPDOLK 890
QY 988 EIG 990
1:1

Db 891 ELG 893

RESULT 6

Q61307 PRELIMINARY: PRT: 1943 AA.

AC Q61307; Q61305; Q61306; Q61308; Q61309; Q61310; Q08866; Q08867;

DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Ankyrin 3 (Ankyrin G) (Epithelial ankyrin) (Ankyrin-3).

OS ANK3.

GN Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

SEQUENCE FROM N.A. (ISOFORMS 1-6).

RP STRAIN=C57BL/6J; TISSUE=KIDNEY;

RX MEDLINE=95340633; PubMed=7615634;

RA Peters L.L., John K.M., Lu F.M., Eicher E.M., Higgins A., Yialamas M.,

RA Turtzo L.C., Otsuka A.J., Lux S.E.;

RT "Ank3 (epithelial ankyrin), a widely distributed new member of the

RT ankyrin gene family and the major ankyrin in kidney, is expressed in

RT alternatively spliced forms, including forms that lack the repeat

RT domain.";

RL J. Cell Biol. 130:313-330(1995).

RN [2]

SEQUENCE FROM N.A. (ISOFORMS 7 AND 8).

RP STRAIN=C3H/HEJ; TISSUE=BONE MARROW;

RX MEDLINE=97213781; PubMed=9060470;

RA Hooch T.C., Peters L.L., Lux S.E.;

RA "Isoforms of ankyrin-3 that lack the NH2-terminal repeats associate

RT with mouse macrophage lysosomes.";

RL J. Cell Biol. 136:1059-1070(1997).

CC -!- FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-

CC SKELETAL ELEMENTS; THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN

CC BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85,

CC AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND

CC DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO

CC THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;

CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.

CC -!- SUBCELLULAR LOCATION: PLASMA-MEMBRANE ASSOCIATED. ISOFORMS 7 AND 8

CC ARE ASSOCIATED WITH THE LYSOSOMAL MEMBRANE.

CC -!- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS; 1 (SHOWN HERE), 2, 3,

CC 4, 5, 6, 7 AND 8; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -!- TISSUE SPECIFICITY: MAJOR FORM OF ANKYRIN IN KIDNEY AND OTHER

CC EPITHELIAL TISSUES; ALSO EXPRESSED IN MACROPHAGES, MEGAKARYOCYTES,

CC LEYDIG CELLS, CARDIAC, SMOOTH AND SKELETAL MUSCLE, INITIAL

CC SEGMENTS OF AXONS, AND NODES OF RANVIER.

CC -!- SIMILARITY: CONTAINS 24 ANK REPEATS.

DR EMBL; L40631; AAB01603.1; -

DR EMBL; L40632; AAB01605.1; -

DR EMBL; U92275; AAB58381.1; -

DR EMBL; L40631; AAB01602.1; -

DR EMBL; L40632; AAB01604.1; -

DR EMBL; L40632; AAB01606.1; -

DR EMBL; L40632; AAB01607.1; -

DR EMBL; U92274; AAB58380.1; -

DR HSSP; P55273; 1B18.

DR MGD; MGI:88026; Ank3.

DR InterPro; IPR002110; ANK.

DR InterPro; IPR000488; Death.

DR InterPro; IPR000906; ZU5.

DR Pfam; PF00023; ank; 24.

DR Pfam; PF00531; death; 1.

DR Pfam; PF00791; ZU5; 1.

DR PRINTS; PR01415; ANKYRIN.

DR SMART; SM00248; ANK; 21.

DR SMART; SM00005; DEATH; 1.

DR SMART; SM00218; ZU5; 1.

DR PROSITE; PS50088; ANK_REPEAT; 21.

DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.

DR PROSITE; PS50017; DEATH_DOMAIN; 1.

DR KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;

FT Multigene family. 826 ANION EXCHANGE PROTEIN BINDING DOMAIN.

FT DOMAIN 23 810 24 X ANK MOTIF REPEATS.

FT DOMAIN 856 1437 SPECTRIN BINDING DOMAIN.

FT DOMAIN 1438 1943 REGULATORY DOMAIN

FT (REGULATES BINDING OF ANKYRIN TO SPECTRIN

FT AND THE BAND 3 PROTEIN).

FT REPEAT 23 55 ANK MOTIF 1.

FT REPEAT 56 88 ANK MOTIF 2.

FT REPEAT 89 121 ANK MOTIF 3.

FT REPEAT 122 154 ANK MOTIF 4.

FT REPEAT 155 183 ANK MOTIF 5.

FT REPEAT 184 216 ANK MOTIF 6.

FT REPEAT 217 249 ANK MOTIF 7.

FT REPEAT 250 282 ANK MOTIF 8.

FT REPEAT 283 315 ANK MOTIF 9.

FT REPEAT 316 348 ANK MOTIF 10.

FT REPEAT 349 381 ANK MOTIF 11.

FT REPEAT 382 414 ANK MOTIF 12.

FT REPEAT 415 447 ANK MOTIF 13.

FT REPEAT 448 480 ANK MOTIF 14.

FT REPEAT 481 513 ANK MOTIF 15.

FT REPEAT 514 546 ANK MOTIF 16.

FT REPEAT 547 579 ANK MOTIF 17.

FT REPEAT 580 612 ANK MOTIF 18.

FT REPEAT 613 645 ANK MOTIF 19.

FT REPEAT 646 678 ANK MOTIF 20.

FT REPEAT 679 711 ANK MOTIF 21.

FT REPEAT 712 744 ANK MOTIF 22.

FT REPEAT 745 777 ANK MOTIF 23.

FT REPEAT 778 810 ANK MOTIF 24.

FT VARSPLIC 1 805 MISSING (IN ISOFORM 7 AND ISOFORM 8).

FT VARSPLIC 1 849 MISSING (IN ISOFORM 2 AND ISOFORM 3).

FT VARSPLIC 833 853 MISSING (IN ISOFORM 7 AND ISOFORM 8).

FT VARSPLIC 834 855 VKRKAPEKLSDGEYISDGEEG -> GDKCTWEKIPKQVEV

FT VARSPLIC 850 855 LVK (IN ISOFORM 5).

FT VARSPLIC 850 855 SDGEEG -> MALPHS (IN ISOFORM 2 AND

FT VARSPLIC 852 852 ISOFORM 3).

FT VARSPLIC 852 852 G -> GDKCTWEKIPKQVEVLVKS (IN ISOFORM 4

FT VARSPLIC 1569 1764 AND ISOFORM 6).

FT VARSPLIC 1943 AA; 212013 MW; B23C4B6F82904804 CRC64; MISSING (IN ISOFORM 3, ISOFORM 4 AND

FT VARSPLIC 1943 AA; 212013 MW; B23C4B6F82904804 CRC64; MISSING (IN ISOFORM 7).

SQ SEQUENCE 1943 AA; 212013 MW; B23C4B6F82904804 CRC64;

Query Match 13.8%; Score 891; DB 11; Length 1943;

Best Local Similarity 28.0%; Pred. No. 1.3e-46;

Matches 339; Conservative 167; Mismatches 495; Indels 209; Gaps 39;

QY 68 AVAAARIMSGRCAGGGAACASAAAEVPAARELFACRNGDVERVKRLVTPKVNRSR 127

DB 10 AKPAHRKRKGGKSDANASYLRAARAGHLEK-----LDYIKNG-----VDVNICN 54

QY 128 TAGRKSTPLHFAAGFGKDVVEYLLONGANVQARDGGLIPLHNACSGHAEVNVLLLRH 187

DB 55 QNGLNA--LHLASKEGHVEVVSSELLQREANVDAATKKGNTALHIASLAGOAEVVKVLVTN 112

QY 188 GADPNARDNNVYTPHLEAAIKGKIDVLCIVLHGAETPIRNTDGTALDLA-----DPSA 242

DB 113 GANVNAQSONGFTPLYNAAQENHLEVVVRLDNGASOSLATEGFTPLVALQOQHVV 172

QY 243 KAVLTGEYK---KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVK 299

DB 173 SLLLENDTKGVRLPALHIAARKDDTKAAALLQNDTNADVESKSGFTPLHIAAHYGNIN 232

QY 300 IVQLLQHGADYHAKDKGLVPLHNACSGHYEYVELLVKHGACVNMADLWQFTPLHEAA 359

DB 233 VATLLNRAAADFTARNDDITPLHVASKRGNANNVKILLDRGAKIDAKTRDGLTPLHCGA 292

QY 360 SKNRVCSLLISYGADPTLLNCHNKSATDLAPTOLKERLAYEFKGHSLIQAAAREADVT 419

DB 293 RSGHEQVVMELDRSA-PILSKTN-----GLSPLHMATQGD-- 328

QY 655 HGAVVNVADLWKFTPLHFAAAGKYEICKLLOHGADPTKKNRDGTPLDL-VKDGDTDI 713
 Db 581 KSASPDAGKSGTLPLHVAHYDNQKVALLLDQASPHAAKNGYPLHHTAAKNQMDI 640
 QY 714 QDLP-----RGDAALDAAKGGCLARVKKLSSPD-NVNCRTDQGRHSTPLHLA 760
 Db 641 ATSLLEYGADANAVTRQGIASVHLAAQEGHVDVMSLLSRANVNLSNKRGLN--PLHLG 698
 QY 761 AGYNNLEVAEYLLQHGADVNAQDKGLPLHNAASYGHVDVAAALLIKYNACVNATDKWAF 820
 Db 699 QGEDRVNVAEYLVNQGAHVDAQTKMGYTPLVHVGCHYGNIKIVNLFQHSKYNATKNGY 758
 QY 821 TPLHFAAQKRTQCALILLAHGADPTLKNQEGQTPDLIVS-----ADDVSALITAAAMP 873
 Db 759 TPLHQAQQGHTHIINVLQNNASPNELTVNGNTALALARRLGYISVVDTLKVVTTEEIMT 818
 QY 874 PSALPSCYK---POVLNGVRSFGATADALSSGSPSSLSAASSLDNLSFSSELSVVVS 930
 Db 819 TTTITERKHKMNPETMNEVLDD--MSDDEVGKG-SAPEKLSHGEXI----- 860
 QY 931 SSGTEGASSLEKKEVPGVDFSIOTFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIG 990
 Db 861 SDGEEG-----EDAITGDTOKSLGPDQLKEIG 887

RESULT 8
 070511 PRELIMINARY; PRT; 2622 AA.

ID 070511; AC 070511; DT 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1999 (Tremblrel. 11, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE 270 kDa ANKYRIN G isoform (ANKYRING) (Fragment).
 GN ANK3
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=98417633; PubMed=9744885;
 RA Zhang X., Bennett V.;
 RT "Restriction of 480/270-kD ankyrin G to axon proximal segments
 requires multiple ankyrin G-specific domains.";
 RL J. Cell Biol. 142:1571-1581(1998).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP Carpenter S.S., Zhang X.;
 RA Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE OF 934-1220 FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SKELETAL MUSCLE;
 RX MEDLINE=98330444; PubMed=9664041;
 RA Kordeli E., Ludosky M.A., Deprette C., Frappier T., Cartaud J.;
 RT "Ankyrin G is associated with the postsynaptic membrane and the
 sarcoplasmic reticulum in the skeletal muscle fibre.";
 RL J. Cell Sci. 111:2197-2207(1998).
 DR EMBL; AF102552; AAC78143.1; -;
 DR EMBL; AF065150; AAC18853.1; -;
 DR HSSP; P55273; IB18.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR004488; Death.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00023; ank; 24.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00791; ZU5; 1.
 DR SMART; SM00248; ANK; 21.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 21.
 DR PROSITE; PS50297; ANK_REP_REGION; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 KW ANK repeat; Repeat.

FT NON_TER 2622 2622
 SQ SEQUENCE 2622 AA; 284456 MW; 67B34830D3AC884E CRC64;
 Query Match 13.7%; Score 885.5; DB 11; Length 2622;
 Best Local Similarity 30.5%; Pred. No. 4.6e-46;
 Matches 283; Conservative 124; Mismatches 373; Indels 147; Gaps 24;
 QY 105 ACRNGDVERVKRLVTPKVNSRDTAGK-STPLHFAAGFRKDVVYLLONGANVQARD 163
 Db 81 ASKEGHEVVSSELLQREA--NVDAATKGTALHIALAGAEVVKVLYVINGANVNAQS 138
 QY 164 GGLIPLHNACSGHAEVNNLLLRHGADPNARNWNVYPLHEAAIKGIDVICVILLQGA 223
 Db 139 NGFTPLYMAAQENHLEVRFLDNGASQSLATEDGETPLAVALQQGHQDVVSLLENDTK 198
 QY 224 PIRNTDGTALDAPSAKAVLTGEYKDELLESARSNGNEEKMMALLTPLNVNCHASDG 283
 Db 199 GVRPLPALHIAAKDDTKAAALL-----LQNDTNDIESKVV-----VNRATESG 243
 QY 284 RKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLVPLHNACSYGHEVTELVKHCAC 343
 Db 244 --FTSLHIAAHYGNINVTALLNRAAAVDFTARNITPLHVASKRGNANVMKLLDRGAK 301
 QY 344 VNAMDLMQFTPLHEAASKNRVCSLLSYGADPTLLNCHNKSALIDAPTQLKERLAYE 403
 Db 302 IDAKTRDGLTPLHCGARSGHEQVEMLLDRAA-PILSKTKN----- 341
 QY 404 FKGSLLQAAAREADVTRIKKHLDSLEVMNFKH-----PQTHETALHCAASPYPKRKQI 456
 Db 342 --GLSPLHMATQGD-----HLNCVQLLLQHNVPDVTNDYLTALHVAHCHGYK--V 390
 QY 457 CELLRKGANINEKTRFTPLHVASEKAHNDVVVVVVKHEAKVNALDNLGQTSUHRAA 516
 Db 391 AKVLLDKKANPAKALNGFTPLHIAKKNRIRVMEILLKHGASIQAVTESGLTPIHVAAF 450
 QY 517 CHLQTCRLLSYGCDPNIIISLOGFTALQM-----NENVOOLLQEGISLNSADRO-- 569
 Db 451 MGHVNTVSQMLMHGASPNNTNVRGETALHMAARSGQAEVRYLVQDGAQVEAKAKDDQTP 510
 QY 570 LLEAAKAGDVTYVKLCITV--QSVNCRDIEGRSTPLHFAAGYNRVSVVVEYLLQHGADVHA 628
 Db 511 LHISARLGRADIVQOLLQOGASPNATTSQ--YTPLHSAREGHEDVAFAFLDHGASLSI 568
 QY 629 KDKGGLVPLHNACSYGHEVAELLVKGAVNVNADLWKFTPLHFAAAGKYEICKLLOH 688
 Db 569 TTKKGFTPLHVAAYKGLVASLLOKASPDAAKSGSLTPLHVAHYDNQKVALLLDQ 628
 QY 689 GADPTKKNRDGTPLDL-VKDGDTDIQDLL-----RGDAALLDAAKKGCLARV 735
 Db 629 GASPHAAAKNGYTPHIAAKKNQMDIATSLLEYGADANVPVTRQGIASVHLAAQEGHVDV 688
 QY 736 KKLSSPD-NVNCRTDQGRHSTPLHFAAGYNVNLEVAEYLLQHGADVNAQDKGLIPLHNA 794
 Db 689 SLLSRANVNLSNKG--LTPLHLAAQEDRVNVAEVLVNOGAHVDAQTKMGYTPLVHVC 746
 QY 795 SYGHVDVAAALLIKYNACVNATDKWAFPLHEAAQKRTQCALILLAHGADPTLKNQEG 854
 Db 747 HYGNIKIVNLFQHSKYNATKNGYTPHLHQAQOQGHTHIINVLQNNASPNELTVNGNT 806
 QY 855 PLDLVS-----ADDVSALLTAAMPSPALPSCYK---POVLNGVRSFGATADALSSG 904
 Db 807 ALAIARRLGYISVVDTLKVVTTEEIMTTTTITEKHKMNPETMNEVLDD-----M 854
 QY 905 SPSSLSAASLDNLS-GSFSSELSVSSSGTEGASSLEKKEVPGVDFSIOTFVRNLGLEH 963
 Db 855 SDDEVGKASAPEKLSDGEY-----ISDGEEG----- 880
 QY 964 LMDIFEREQITLDVLVEMGHKELKEIG 990
 Db 881 -----EDAITGDTOKYLGPQDLKEIG 901

RESULT 9

[illegible]


```
DR PROSITE; PS50088; ANK_REPEAT; 21.  
DR PROSITE; PS50297; ANK_REP_REGION; 1.  
DR PROSITE; PS50017; DEATH_DOMAIN; 1.  
KW ANK repeat; Repeat.  
SQ SEQUENCE 1549 AA; 170167 MW; 3ECE852B1484DFF CRC64;  
  
Query Match 13.4%; Score 866.5; DB 5; Length 1549;  
Best Local Similarity 27.4%; Pred. No. 2.3e-45;  
Matches 311; Conservative 174; Mismatches 491; Indels 159; Gaps 34;  
  
QY 98 AARELFACRNGDVERVKRLV---TPEKVNSEDRTAGRKSTPLHFAAGFGRKDVVEYLQ 154  
DB 38 ATISFLRAARSGDIKKVWDFDCGGEISDINSCNANGLNA--LHLAAKDGYYDICCELLR 95  
  
QY 155 GANYQARDGGGLIPLHNACSFGEHAEVNNLLRHGADPNARDNNVYTLPEHAAIKGIDVC 214  
DB 96 GIKIDNATKNGTALHIAASLAGQHDVINOILLYNANVNVQSLNGFTPLYMAAQENHDNCC 155  
  
QY 215 IVLQHGAEPTIRNVDGRTALDLA-----DPSAKAVLGEYK---KDELLESARSGNEEK 266  
DB 156 RTLLANGANPSLSTEDGFTPLAVAMQOQHKDITVAVLLENDVRGKVRPLPALHIAAKKNDVN 215  
  
QY 267 MMALLTPLNVNCHASDGRKSTPLHAAACYNRVKIVOLLQHGADVHAKDKGLVPLHNAC 326  
DB 216 AAKLLQHDPNADIVKSGFTPLHIAHYGNVDIATLLNKNKADVNYVAKHNIITPLHVAC 275  
  
QY 327 SYGHEVTELLVKGACVYAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN-K 385  
DB 276 KWGKLSLCTLLCRGAKIDAATROGLTPLHCASRSGHVEVTKHLLQNA-PILTKNGL 334  
  
QY 386 SAIDLAPTPQKRLAYEFKGSLLQAAAREADVTRIKKHSLE-----MVNF 432  
DB 335 SALHMAAGE-----HDEAAHLLDNKAPVDEVTVDTALHVAAHGCHVKVAKLLDY 388  
  
QY 433 K-HQOTHE---TALHCAAAAPYKPKQICELLKRGKININEKTEFTPLPHVASEKHN 487  
DB 389 KANFNARALNGFTPLHHTACK---NRKWEELLKHGAGICATTESGLTPLHVASEMGC 445  
  
QY 488 DVEVVKHAEKVALDNLQGTSLHRAAYCGHLQTCRLLLSYCGDPNIIISLQGTAL--- 544  
DB 446 NIVTYLLQHEASADLPTIRGETPLHLAARANAQADIIRILLR-SAKVDAIAREGOTPLHVA 504  
  
QY 545 -QMGNEH-VOLLQEGISLGNSEADR--OLLEAAKAGDVEVTKLCTVQSVNCRDIEGRQ 600  
DB 505 SRLGNINILMLLQHGAEINAQSDKYSALHIAAKEGQENTVQVLLNGAEN-NAVTKKG 563  
  
QY 601 STPLHFAAGYNRVSWVEYLLQHGADVHAKDKGLVPLHNACSYGHEVYAEVLLVKGAVN 660  
DB 564 FTPLHLACKYQKNVQVOLLONGASIDFCQKNDVTPLVHATHYNNPSIVLELLKNGSSPN 623  
  
QY 661 VADLWKFTPLHEAAKGYEICKLLQHGADPTKKNRDGNTPDL-VKDGDTDIQDLRG 719  
DB 624 LCARNGQCAIHIACKKNYLETAMQLLQHGADVNIISKSGFSLHAAOGGNDVMVQLLE 683  
  
QY 720 DAALLDAKKGCLARVKLLSPDNVNCRTQGRHSTPLHAAAGYNNLEVAEYLLQHGADV 779  
DB 684 YGVLSAAKNGKL-----TPLHAAQEGHVLVSQILLEHGANI 720  
  
QY 780 NAQDKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFTPLHEAAAGRTOLCALLL 839  
DB 721 SERTNGYTPLHMAHYGHLDLVKFFENDADTEMSSNIGYTPLVHQAQOQHMIINILL 780  
  
QY 840 AHGADPTLKNOEGOTPLDLYSADVSALLTAAMPSPALPSCYKPVQLNG-----V 889  
DB 781 RHRANPNALTKDGNALHIAASN-----LGYVTVMESLKIIVTSTSVINSIGAIEEKLKV 834  
  
QY 890 RSPCATDALSSGSPSSLSAAS-----SLDNLSGSFSELSVVSSSGTE-----GAS 938  
DB 835 MTELMQETLLSDDBDSCDDLDDHNNHYKYMATDDLKANTGQDQKNPDTTNTDHLTDVS 894  
  
QY 939 SLEKKEVPGVDFSTQVRNMLGLEHMD--IFEREQITLDVLVEMGHGKELKEIGINAYGH 996  
DB 895 VLANKKEILPNEMSCIELTE---IGHKPDVNVIAARSQVHLGFLVSF---LVDARGGSMRGY 948
```

```
QY 997 RHKLKQVE-----RLISQOGLNPYLTLNTISGSGTILLDLSPDDKBFQS--- 1041  
DB 949 RINGVRIIVPPKACAEPTITCRVYKQPVNVNPPPLMEGEALVSRILEMSPYDGMFLSP 1008  
  
QY 1042 -VEEEMQSTVREHRDGGHAGGIFNRYNLIKQKVCNKKLWERYTHRRKRVSEENHNHANE 1100  
DB 1009 TLEVPHYGTLRK-----NEREII-ILRSNGESWREH-NLYKDIIGEDINQTEE 1055  
  
QY 1101 RMLFHGSPFVNAITHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIG--GGT 1153  
DB 1056 ---FHSDDRIV-IVTQNVPHFFAVV-----SRVRQEVHVIGPDGDT 1092  
  
RESULT 13  
Q24241 ID Q24241 PRELIMINARY; PRT; 1549 AA.  
AC Q24241;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE ANKYRIN.  
GN ANK OR CG1651.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A631;  
RX MEDLINE=95024098; PubMed=7937942;  
RA Dubreuil R.R., Yu J.;  
RT "Ankyrin and beta-spectrin accumulate independently of alpha-spectrin in Drosophila.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:10285-10289(1994).  
DR EMBL; L35601; AAC37208.1; --  
DR HSSP; P80144; 2MVO.  
DR FlyBase; FBgn0011747; Ank.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR000906; ZU5.  
DR Pfam; PF00023; ank; 23.  
DR Pfam; PF00531; death; 1.  
DR Pfam; PF00791; ZU5; 1.  
DR PRINTS; PR01415; ANKYRIN.  
DR SMART; SM00248; ANK; 21.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00218; ZU5; 1.  
DR PROSITE; PS50088; ANK_REPEAT; 21.  
DR PROSITE; PS50297; ANK_REP_REGION; 1.  
DR PROSITE; PS50017; DEATH_DOMAIN; 1.  
KW ANK repeat; Repeat.  
SQ SEQUENCE 1549 AA; 170096 MW; 587FB6BF88E657CE CRC64;
```

```
Query Match 13.4%; Score 866.5; DB 5; Length 1549;  
Best Local Similarity 27.4%; Pred. No. 3.1e-45;  
Matches 311; Conservative 173; Mismatches 492; Indels 159; Gaps 34;
```

```
QY 98 AARELFACRNGDVERVKRLV---TPEKVNSEDRTAGRKSTPLHFAAGFGRKDVVEYLQ 154  
DB 38 ATISFLRAARSGDIKKVWDFDCGGEISDINSCNANGLNA--LHLAAKDGYYDICCELLR 95  
  
QY 155 GANYQARDGGGLIPLHNACSFGEHAEVNNLLRHGADPNARDNNVYTLPEHAAIKGIDVC 214  
DB 96 GIKIDNATKNGTALHIAASLAGQHDVINOILLYNANVNVQSLNGFTPLYMAAQENHDNCC 155  
  
QY 215 IVLQHGAEPTIRNVDGRTALDLA-----DPSAKAVLGEYK---KDELLESARSGNEEK 266  
DB 156 RTLLANGANPSLSTEDGFTPLAVAMQOQHKDITVAVLLENDVRGKVRPLPALHIAAKKNDVN 215  
  
QY 267 MMALLTPLNVNCHASDGRKSTPLHAAACYNRVKIVOLLQHGADVHAKDKGLVPLHNAC 326
```



```
Db 640 TDMVTLVLEKGANIHMSTKSGTSLTSLAAAEKVNVDILTKHGADQDAYTKLGYTPLLIV 699
Qy 859 V-----SADOVSAALLTAAMPSSA-LPSCYKP-----QVLNGVRSFGATADALSSGP 903
Db 700 ACHYGKVNKVNELLKOGANVNKATKNGYTPHLHQAQQGHTHTIINVLLQHCAPNATTA-- 757
Qy 904 SSPSLSSAASSLDNLGSGFSFSSVSSSGTEGASLEKKEKPGVDFTSFQFVRNLGLEH 963
Db 758 NGNTALAIKRLGYI--SVVDTLKVVTEVTITITTEKHKLNAPEITMIE----- 806
Qy 964 LMDIFERE---OITLDVLVEMGHKELKEIG 990
Db 807 VLVDSDDEGGDDVTGDDGLEYLRPDLKELG 836

RESULT 15
Q17487 PRELIMINARY; PRT; 1809 AA.
AC Q17487;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE E. elegans ankyrin-related unc-44 (GB:U21734).
GN UNC-44.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Smaildon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinon-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL elegans.";
RN Nature 368:32-38(1994).
[2]
RN RP SEQUENCE FROM N.A.
RA Gattung S.;
RT "The sequence of C. elegans cosmid B0350.";
RN Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
[3]
RN RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U50071; AAA93446.1; -.
DR HSSP; P42773; 11HB.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR001360; GH_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 22.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS500572; GLYCOSTYL_HYDROL_F1_1; UNKNOWN_2.
KW ANK repeat; Repeat.
SQ SEQUENCE 1809 AA; 198193 MW; AD86349AA64534F8 CRC64;
```

```
Query Match 13.3%; Score 858; DB 5; Length 1809;
Best Local Similarity 26.3%; Pred. No. 1.3e-44;
Matches 319; Conservative 163; Mismatches 443; Indels 286; Gaps 39;

Qy 91 AAFAVEPAARE---LFEACRNGDVERVRVLY--TPKVNSRDTAGRKSTPLHFAAGFGRK 145
Db 22 AAPAPEPGRAEGSASFRAARAAGDLEKLELLRAGTDINTSNAGLNS--LHLASKEGHS 79
Qy 146 DYVEYLLQGANVQARDGGCLPLHNACSGHAEVNVNLLRHGADPNARDNNNYTPHHEA 205
Db 80 EVVRELKFAQOVDATRGNTALTALHTASLAGOSLITVLVENGANVVOVGFTPLYMA 139
Qy 206 ATKGIDVCIVLQHGAEPTINTDGTALDLADPSAKAVLTGEYKDELLESARSNEE 265
Db 140 AQENHEEVVYKLLKHGANQALSTEDGFTPLAVA-----LQOGRDR 179
Qy 266 KMWALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLQHGADVHAHKGDLVPLHN 324
Db 180 VYAVLLE-----NDSKGVRLPALHIAAKKDDTTAATLLQLQNEHNPVDTSGFTPLHI 233
Qy 325 ACSYGHYEVELLVKHGACVNMALWQFTPLHEAAKKNRVEVCSLLLSYGADPTLLNCHN 384
Db 234 AAHYGHENVGQLLEKGANVYQARHNISPLHVATKWGTNNANLLLSRGA---IIDSRT 290
Qy 385 KSAIDLAPTPOLKERLAYEFKCHSLLOQAAREADVTRIKKHLSELMVNFKHPOTHTALHC 444
Db 291 KDLI-----TPLHC 299
Qy 445 AAAPYKRKQICELLRRKANINEKTEFLTPLHVAASEKAHNDVVEVVVHKAQVN--A 502
Db 300 AARSQH---DQVVDLLVVGAPISAKTKNGLAPLHMAAGDHDVAARTLLYHRAPVDDVT 356
Qy 503 LDNLGQTSLHRAAYCGHLOTCLLSYGCDPNIISLOGFTALQMGNE-----550
Db 357 VDYL--TPLHVAACHGVRAKLLDRSADPNRNLNGFTPLHITACKKRIKVVVLLKY 414
Qy 551 ---VQQLQOEGIS-----LGNSEADRLLEAKAGDVEIVKCLKTVOVNCRDIEGRQ 600
Db 415 RAAEATESTESGTPHVAAFMGAINIVYLQOAGNPVDET-----RG 458
Qy 601 STPLHFAAGYNRVSVVEYLQHGADVHAHKGDLVPLHNACSYGHEVAELLVKGAVVN 660
Db 459 ETPLHFAARANTDVRVRLIRNGAKVDAQARELOTPLHIAKRLGNTDIVILLQGANSN 518
Qy 661 VADLWKEFTPLHEAAKGYEICKLLLOHGADPTKKNRDGNTPLDLV-KDGTDTIQDLL-- 717
Db 519 ATTRDNYSPHLHAAKEGEEVAGILLDHNADKTLTKKGFPLHASKYGNLEVVRLLE 578
Qy 718 RGDAAALLDAARKGCLARKVKKLSSPDNVNCRDTCGRHS--TPLHLAGYNNLEVAEYLLQHG 776
Db 579 RGTVP-----DIEGKNQVTPHVAAHYNNNDKVMALLJENG 613
Qy 777 ADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFPTPLHEAAKGTQLCA 836
Db 614 ASAKAAAKNGYTPHLHIAAKKNQMEIATLLQFKADPNAKSRAGFTPLHLSAQEGHKEISG 673
Qy 837 LLLAHGADPTLKNOEGOTPLDLYSADD-----VSALLTAAMPSPSALPSCYK 882
Db 674 LLTENGSDVGAKANNGGLTAMHLCAQEDHPVPAQILYNGAEINSKTNAGYTPLHV-ACHF 732
Qy 883 POVLNGVR---SPCATADALSSGSPSSLSAASSLSGSELSVSVSSSGTEGASS 939
Db 733 GQ-LNMVKFLVENGADVGEKTRASYTPHLHQAQQGHNVCVRYLLE-----NGASP 781
Qy 940 LEKKEVPGVDFSTQFVRNLG-----LEHLMDFEREQITLDVLVEMGHKELKEIGIN-- 992
Db 782 NEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTVDERYKQPNPEAMNEM 838
Qy 993 -----AYGHRHKL-----KGVRLISGOOGLNPYLTNTSG---1024
Db 839 FSESEDEGQAHEVAHAHEKDFSDNLQTGLQDSTGVHMTHTGEQLQSRQSELENGGAIP 898
```

Wed Feb 12 16:08:54 2003

QY 1025 ----SCTILIDLSPPDKFEQSVEEEMQS-----TVREHRDG--CHAGGITFNRYNLIKOKVC 1075
Db 899 KINSQ---GNSP-EKEFAKIAPVATSSPIATSNSSQSGIAPRAGSISGQFQQOPLHGAG 953
QY 1076 NKKLWERYTHRKE--VSEENHNHANERMLFHS-----PFVNAIIHKGf-----DERHA 1123
Db 954 PEDNLEELVRRQNHPIINAGNYDNGGVAMLENGHADNVPIGHHVTQPSFLISFLVDARGG 1013
QY 1124 YIGGMFGAGIYFAENSXSNQYVYIGGGTGCPCVHKDRSCYICHRQLLFCRVTLGKSFLO 1183
Db 1014 AMRGCHSGVRIIVPPRKASQ-----PIR-----VTCRY-LRKD--- 1046
QY 1184 FSAMKMAHSPP 1194
Db 1047 ----KLAHPPP 1053

Search completed: February 12, 2003, 07:44:28
Job time : 50.6325 secs

